

Tabular Reporting with R

Charles Roosen, PhD
Technical Director
Mango Solutions AG

October 26, 2011



Agenda

- Overview
- Latex Documents
 - Generating Tables: xtable
 - Inserting Tables: `\input` or Sweave
- Microsoft Word Documents
 - Generating Tables: Mango reporting library
 - Inserting Tables: Push2Doc

Overview

- Why create tables?
 - Provide data listings or summaries
- Where do tables go?
 - Latex documents
 - Word documents
- What makes up a table?
 - Data values
 - Column labels, row labels, row groupings
 - Styling: Borders, shading, column widths
 - Captions

LATEX

Generating Tables: xtable

- Widely used library for Latex or HTML table generation
 - Created by David Dahl
 - Now maintained by Charles Roosen
 - On CRAN with development hosted on R-Forge

<http://r-forge.r-project.org/projects/xtable/>

Creating an xtable object

- Use the “xtable()” function to create an “xtable” object
 - Generic function with methods for many object types
 - Includes methods for: `anova`, `aov`, `aovlist`, `coxph`, `data.frame`, `glm`, `lm`, `matrix`, `prcomp`, `summary.aov`, `summary.aovlist`, `summary.glm`, `summary.lm`, `summary.prcomp`, `table`, `ts`, `zoo`
 - Other packages implement methods for “xtable()”
- Object has properties regarding formatting of numeric values
 - Utility functions available to get/set these properties

```
> objects("package:xtable")
```

```
[1] "align"      "align<-"    "caption"    "caption<-"
```

```
[5] "digits"     "digits<-"  "display"    "display<-"
```

```
[9] "label"      "label<-"   "xtable"
```

Printing an xtable object

- The “print()” method for “xtable” generates Latex or HTML
- Lots of arguments controlling the markup used

```
print(x, type="latex", file="", append=FALSE, floating=TRUE,  
      floating.environment="table", table.placement = "ht",  
      caption.placement="bottom",  
      latex.environments=c("center"), tabular.environment = "tabular",  
      size=NULL, hline.after=c(-1,0,nrow(x)), NA.string = "",  
      include.rownames=TRUE, include.colnames=TRUE, only.contents=FALSE,  
      add.to.row=NULL, sanitize.text.function=NULL,  
      sanitize.rownames.function=sanitize.text.function,  
      sanitize.colnames.function=sanitize.text.function,  
      math.style.negative=FALSE, html.table.attributes="border=1", ...)
```

Example: Data Frame

```
> data(tli)
> tli.table <- xtable(tli[1:10, ])
> digits(tli.table)[c(2, 6)] <- 0

> print(tli.table, floating = FALSE)
```

	grade	sex	disadv	ethnicity	timth
1	6	M	YES	HISPANIC	43
2	7	M	NO	BLACK	88
3	5	F	YES	HISPANIC	34
4	3	M	YES	HISPANIC	65
5	8	M	YES	WHITE	75
6	5	M	NO	BLACK	74
7	8	F	YES	HISPANIC	72
8	4	M	YES	BLACK	79
9	6	M	NO	WHITE	88
10	7	M	YES	HISPANIC	87

Example: aov

```
> fm1 <- aov(tlimth ~ sex + ethnicity + grade + disadv, data = tli)
> fm1.table <- xtable(fm1)
> print(fm1.table, floating = FALSE)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
sex	1	75.37	75.37	0.38	0.5417
ethnicity	3	2572.15	857.38	4.27	0.0072
grade	1	36.31	36.31	0.18	0.6717
disadv	1	59.30	59.30	0.30	0.5882
Residuals	93	18682.87	200.89		

Example: time series

```
> temp.ts <- ts(cumsum(1 + round(rnorm(100), 0)), start = c(1954,  
+ 7), frequency = 12)  
> temp.table <- xtable(temp.ts, digits = 0)  
> caption(temp.table) <- "Time series example"  
  
> print(temp.table, floating = FALSE)
```

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1954							-1	1	1	2	2	3
1955	6	8	10	11	12	13	13	14	16	16	17	19
1956	20	21	22	21	22	21	22	21	22	23	23	25
1957	26	26	27	27	28	28	29	31	32	34	34	34
1958	35	36	38	40	42	43	42	43	46	47	48	49
1959	49	50	50	51	53	53	56	58	58	59	60	60
1960	62	62	64	66	68	69	70	71	71	71	74	75
1961	75	75	76	77	78	80	82	83	85	87	88	88
1962	88	88	86	87	88	89	90	90	93	94		

Inserting Tables in Latex

- Inserting tables with “\input”
 - Step 1: Write the Latex to a file
 - Step 2: Reference this from the main Latex document with “\input”
- Inserting tables with “Sweave”
 - Step 1: Write an “Rnw” file which is Latex with R code blocks
 - Step 2: Run “Sweave” to generate a Latex document with the R code evaluated and results inserted

MICROSOFT WORD

Using Word for Reports

- Microsoft Word has many features good for collaborative editing
 - Widely used
 - Spelling and grammar checkers
 - Review change tracking

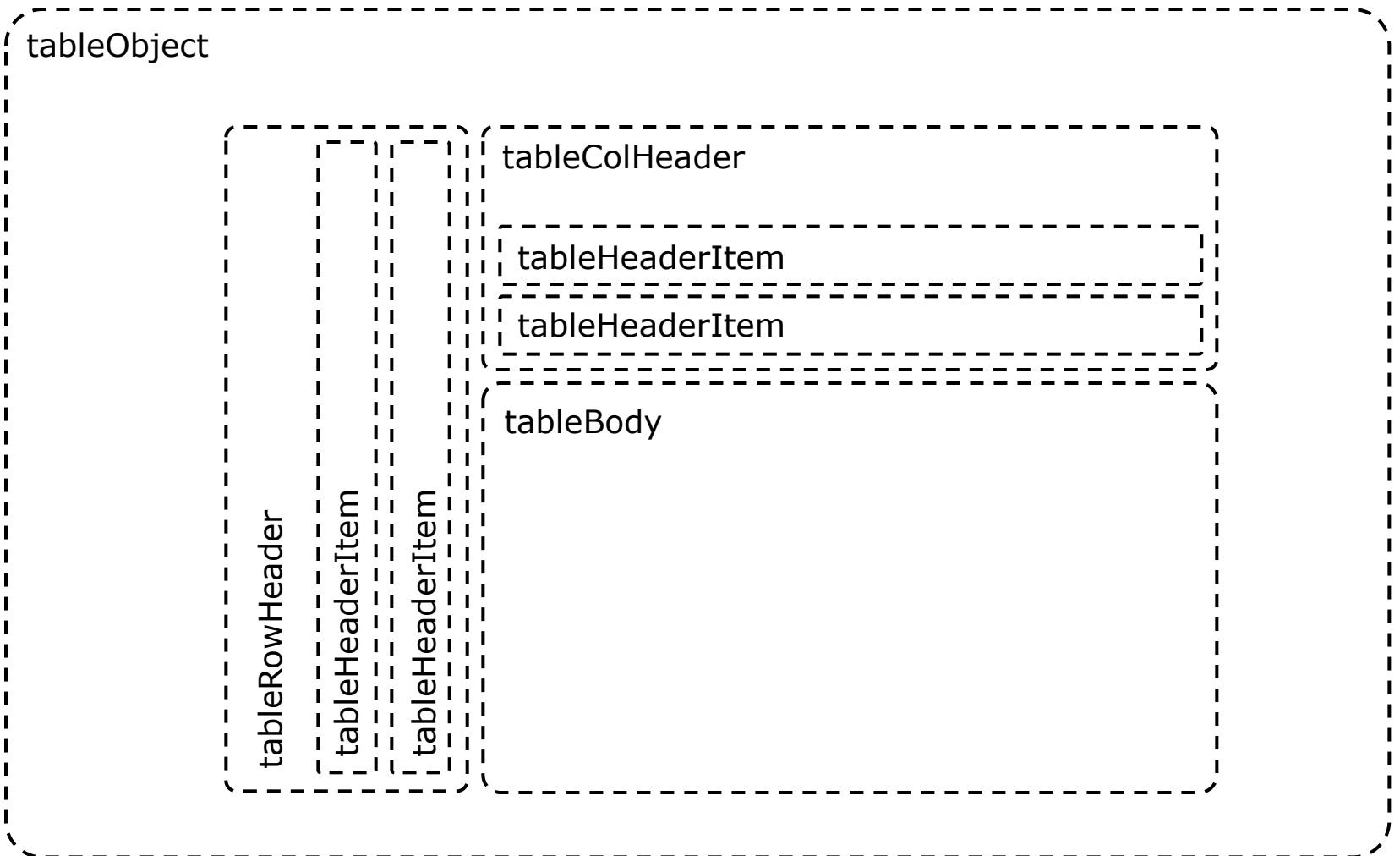
Structuring the Work

- Sweave is good for many uses, but ...
 - Structure is one text file mixing Latex and R
- There are benefits to separating the narrative text from the code
 - Easier to read and review
 - Non-programmers won't understand the code
 - No “compile” cycle needed to see the updates
- One approach
 - Write the narrative of the text in Word
 - Include references to tables and figures
 - Have separate script to generate the tables and figures
 - Use a tool to insert the figures into the document

- Creates PDF, RTF, and HTML reports containing tables and figures
- Rich set of table formatting capabilities

Sex	Dose	Observed Response				
		Mean	Median	Confidence Interval		N
				Lower	Upper	
Male	0	53.6	54.2	40.6	60.1	25
	15	72.9	75.0	62.7	89.8	25
	30	89.0	89.6	67.9	98.3	23
Female	0	49.3	50.8	38.4	78.0	20
	15	65.7	68.2	55.2	82.4	20
	30	78.1	79.9	58.1	86.1	19

Table Objects



R Functions

```
> # Initialize Report (Environment)
> createReport()
> # Add a lattice plot
> addGraphItem( xyplot( Y ~ X, data = myData ) )
> # Add a Summary Table
> addTableItem( summaryTable(myData, c("X", "Y")) ) )
> # Add a Data Table (basically a table of data)
> addTableItem( dataTable( head(myData) ) )
> # Add some text
> addTextItem( LETTERS )
> # Add a small histogram
> addGraphItem( hist(rnorm(100)) )
> # Export this Environment to a document
> publishReport("output.doc")
```

How it's done

- Architecture of the Reporting Library
 - R generates XML
 - Java code reads the XML and creates “iText” objects
 - “iText” used to create the PDF, RTF, HTML documents
- Usage in the Reporting Workflow
 - Can create a complete document, such as a study report appendix
 - Can create a document with a single table for inclusion within a study report

- push2doc is a tool which allows users to insert pre-existing report items (graphs, images, tables and text) into a Word document.
- Report items can be generated by the user's tool or language of choice enabling push2doc to integrate seamlessly into the user's workflow.
- push2doc is a command line application which can be run on Windows via a batch file and on Linux via a shell script.
- Additionally, it can be invoked from another modelling language e.g. R, SAS, Matlab.
- An R wrapper is included in the distribution which allows push2doc to be called directly from R.

Push2Doc Usage

- Push2Doc Steps
 - Create table and figure files for insertion
 - Create a Word (docx) template file
 - Call Push2Doc to create the Word document with insertions
- Platforms
 - Cross-platform: Windows, Linux, Mac, etc.
 - Implemented in Java with no Microsoft Office dependencies
- Execution
 - Simply call Java from a batch file (*.bat) or shell script (*.sh)
 - Self contained set of tools with no install other than Java

Table Specification

- Table location and settings are specified by creating a “Comment” field in the Word document
 - Insert Menu -> Quick Parts -> Field... -> select ‘Comments’ from Field names list.
 - In the textfield marked ‘New comment’ add bookmark using the following syntax:
modreport:BookmarkIdentifier[local | list of options]
- Table data is in a CSV file
 - Can include notation for superscripts, subscripts, Greek symbols
 - Can specify hierarchical tables using a column of row “roles”

Example: Simple Table

Fruit,Quantity,Cost

Apples

Gala,10,£5

Cox,15,£7.50

Discovery,10,£4.55

Oranges

Seville,10,£5

Valencia,15,£7.50

Could write totals in the footer

Fruit	Quantity	Cost
Apples		
Gala	10	£5
Cox	15	£7.50
Discovery	10	£4.55
Oranges		
Seville	10	£5
Valencia	15	£7.50
Could write totals in the footer		

Example: Hierarchical Table

```
head,Fruit,Quantity,Cost
subheader,Apples
datarow,Gala,10,£5
datarow,Cox,15,£7.50
datarow,Discovery,10,£4.55
subheader,Oranges
datarow,Seville,10,£5
datarow,Valencia,15,£7.50
footer,Could write totals in the footer
```

```
modreport;Tag1[local|title=Plot of DV vs PRED|image=Image1.png|caption=Figure Caption]
```

Fruit	Quantity	Cost
<i>Apples</i>		
<i>Gala</i>	<i>10</i>	<i>\$5</i>
<i>Cox</i>	<i>15</i>	<i>\$7.50</i>
<i>Discovery</i>	<i>10</i>	<i>\$4.55</i>
<i>Oranges</i>		
<i>Seville</i>	<i>10</i>	<i>\$5</i>
<i>Valencia</i>	<i>15</i>	<i>\$7.50</i>
<i>Could write totals in the footer</i>		

Summary

- Latex tables and documents
 - xtable to create tables
 - Sweave or `\input` to insert tables
- Word tables and documents
 - Mango reporting library for PDF, RTF, HTML
 - Mango Push2Doc for insertion of tables/figures into Word

Contact Us

Mango Solutions AG

Aeschenvorstadt 36

4051 Basel

Switzerland

+41 (0) 61 20692 92

croosen@mango-solutions.com