

# Export Pivot Table to R Using RExcel

Keith Halbert<sup>1</sup>   Richard Heiberger<sup>1</sup>   Erich Neuwirth<sup>2</sup>

<sup>1</sup>Department of Statistics, Fox School of Business  
Temple University

<sup>2</sup>Didactic Center for Computer Science and Institute for Scientific Computing  
University of Vienna

useR! 2010  
Gaithersburg, MD

# Contents

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

- 1 About RExcel
- 2 About Pivot Tables
- 3 Why Transfer to R?
- 4 How it Works
- 5 Example
- 6 Conclusion
- 7 Appendix

# RExcel

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

Microsoft Excel, available on almost everyone's machine, provides a familiar interface with which users are comfortable. RExcel is a free add-in to Excel that places the full power of R at the Excel user's fingertips.

- provides menu access to many R functions directly from Excel by placing the Rcmdr menu on the Excel menu bar
- capable of data transfer to and from R
- allows the placement of R graphs into the Excel automatic recalculation model
- allows *any* R function to be used in Excel formulas

# RExcel

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

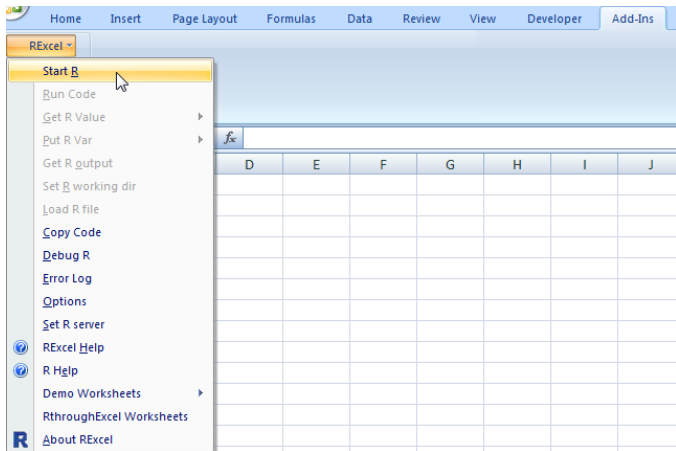
How it Works

Example

Conclusion

Appendix

Once installed, the user connects Excel to R through the Add-ins menu. The help documentation can be accessed from here as well.



# Pivot Tables

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

Pivot tables are data summarization tools common to spreadsheet software (such as Microsoft Excel). They provide a means for quick and intuitive tabulation of data.

The simple data set and pivot table on the right demonstrates this capability. Here, the average of Height is shown with Sex as a row variable.

	A	B	C
1	Name	Sex	Height (in)
2	Billy	Male	70
3	Herbie	Male	64
4	Steve	Male	68
5	Monica	Female	60
6	Kathy	Female	61
7	Susan	Female	65

Row Labels	Average of Height (in)
Female	62.0
Male	67.3
<b>Grand Total</b>	<b>64.7</b>

# Pivot Tables

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

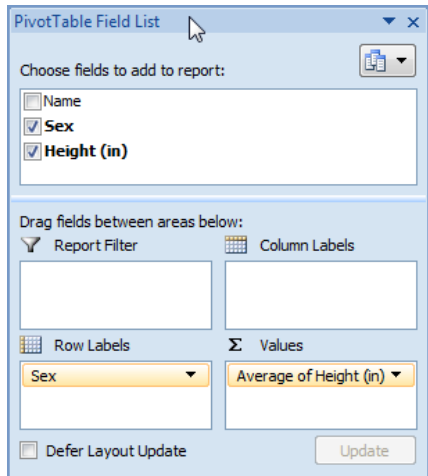
Conclusion

Appendix

The variable selection is made through use of the Field List shown on the right.

Here, Sex has been placed on the rows and we have selected the average of Height to be our summary function.

Excel allows for fields to be dragged, dropped, and filtered in this list.



# When Pivot Tables Are Used

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

- to summarize large transactional data sets
- to find relationships and groupings within data
- to quickly filter data sets
- to organize data in a format that is easy to chart

*Pivot tables are the single most powerful feature in all of Excel.*

– Bill Jelen and Michael Alexander, *Pivot Table Data Crunching*

# Why Transfer a Pivot Table to R?

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

- R graphics are fantastic
- R has many powerful tools for manipulating and analyzing multidimensional tables
- re-creation in R of an existing Excel pivot table could introduce error; automatic transfer is safer



# How it Works

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

PutPivottable has been included in RExcel since February 2010.

The table created in R is of class `structable`, a flat contingency table provided in the excellent `vcd` package (Visualizing Categorical Data).

Options for transfer:

- right-click on a pivot table
- use VBA macro `RInterface.PutPivottable` (not discussed in these slides)

# The Right-Clicking Option

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

With RExcel activated, the user right-clicks anywhere in a pivot table and selects Put Pivottable.

Row Labels	Average of Height (in)
Female	62.0
Male	67.3
<b>Grand Total</b>	<b>64.7</b>

The screenshot shows a pivot table with the following data:

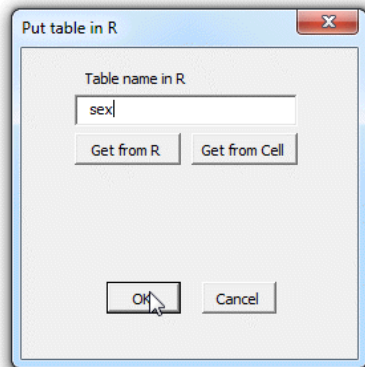
Row Labels	Average of Height (in)
Female	62.0
Male	67.3
<b>Grand Total</b>	<b>64.7</b>

The context menu is open over the Grand Total cell, showing the following options:

- Copy
- Format Cells...
- Number Format...
- Refresh
- Sort
- Move
- Remove "Average of Height (in)"
- Summarize Data By
- Show Details
- Show/Hide Fields
- Value Field Settings...
- PivotTable Options...
- Hide Field List
- Put Pivottable** (highlighted)

# The Right-Clicking Option

The previous action brings up the Put table in R dialog box, where the name is assigned. After execution the table may be manipulated as desired using RExcel or the R console.



# A Classic Example

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

The titanic data consist of the following variables regarding the passengers of the ill-fated maiden voyage of the Titanic:

- Class: levels first, second, third and crew
- Age: levels child and adult
- Sex: levels M and F
- Survive: levels die and live

These are recoded from the original. See the Reference section for more information. Two of the 2201 records follow.

Class	Age	Sex	Survive
third	child	F	die
crew	adult	M	live

# Titanic Example

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

We begin with a simple statistical question: Does the class of the passenger help to explain survival?

We place `Class` on the rows and `Survive` on the Columns.

Counts	Survive		Grand Total
	die	live	
Class			
crew	673	212	885
first	122	203	325
second	167	118	285
third	528	178	706
Grand Total	1490	711	2201

# Titanic Example

To perform a chi-square test of independence, Excel requires us to manually create a table of expected cell counts. These are the numbers we would expect if Class and Survive were independent.

	A	B	C	D	E	F	G	H
3	Counts	Survive						
4	Class	die	live	Grand Total			die	live
5	crew	673	212	885	crew	=D5*B	285.89	
6	first	122	203	325	first	220.01	104.99	
7	second	167	118	285	second	192.94	92.06	
8	third	528	178	706	third	477.94	228.06	
9	Grand Total	1490	711	2201				

CHITEST() calculates a p-value (which indicates dependence).

	A	B	C	D	E	F
10						
11	p-value	5.00E-41				

# Titanic Example

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

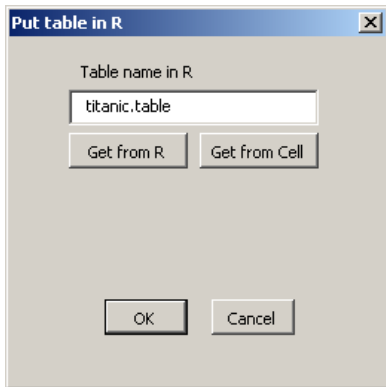
How it Works

**Example**

Conclusion

Appendix

The chi-square test is an approximation. For some count data we may require Fisher's exact test, which Excel does not provide. To perform either of these tests in R, first we transfer the table and give it a sensible name.



# Titanic Example

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

The RExcel function `REval()` evaluates R code and brings the result to Excel. The chi-square approximation is fine for these data. The  $p$ -values are similarly tiny.

In `chisq.test()` R calculates the table of expected values directly from the pivot table. The manual calculation is not necessary.

	result	formula
chi-square in R	5.00E-41	=REval("chisq.test(titanic.table)\$p.value")
Fisher Exact in R	5.29E-39	=REval("fisher.test(titanic.table, workspace=1000000)\$p.value")

Even in this simple example, Excel cannot compete with R when it comes to statistical analysis. We now turn our attention to graphics.



# Titanic Example

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

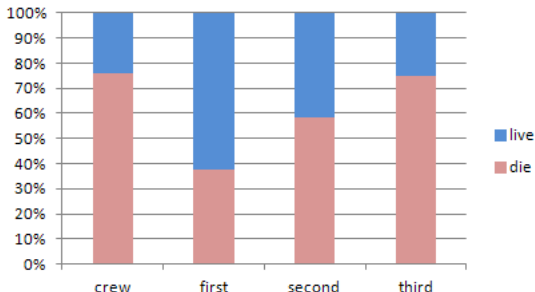
How it Works

Example

Conclusion

Appendix

Below is a standard Excel plot of our pivot table.



This is a simple table and Excel's plot is OK. Next we examine a mosaic plot from package `vcd`. These are great for comparing count data, but are not available in Excel.

# Mosaic Plot

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

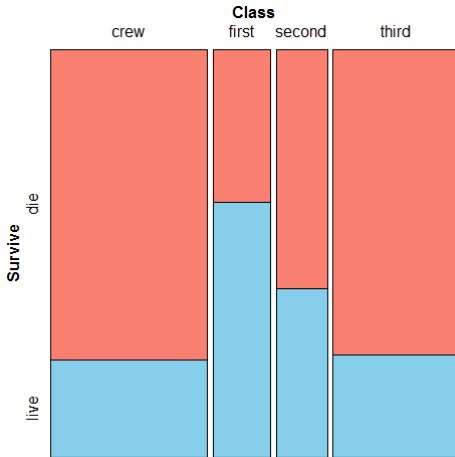
Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix



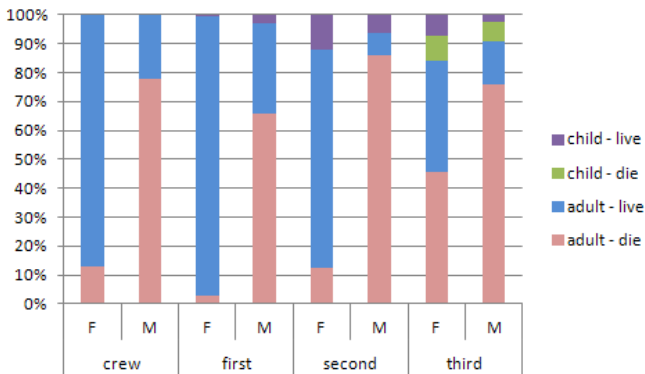
# Titanic Example

We add two additional fields: a second row variable (Sex) and a second column variable (Age).

Count of Survive2		Age	Survive		Grand Total	
Class	Sex	adult	child	die	live	
		die	live			
crew						
	F	3	20	0	0	23
	M	670	192	0	0	862
first						
	F	4	140	0	1	145
	M	118	57	0	5	180
second						
	F	13	80	0	13	106
	M	154	14	0	11	179
third						
	F	89	76	17	14	196
	M	387	75	35	13	510
Grand Total		1438	654	52	57	2201

# Titanic Example

Below is the best we can do in Excel. The mosaic plot on the next slide is superior. Excel pivot table users should have access to these plots.



# The Mosaic Plot is Superior

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

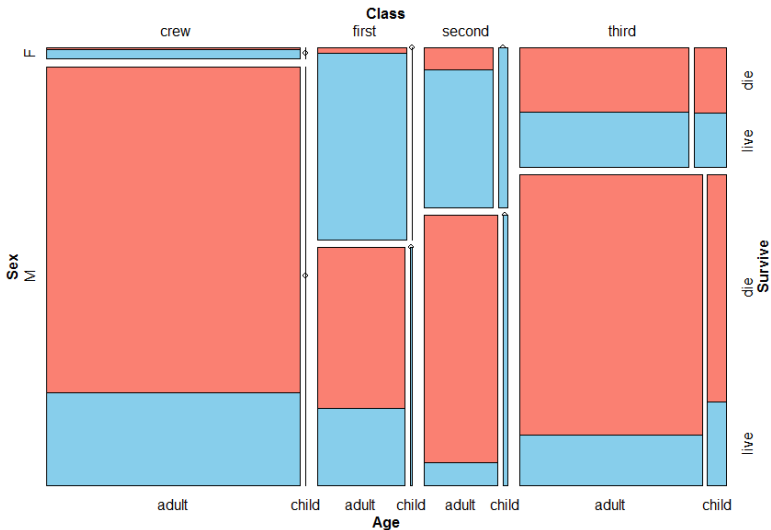
Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix



# Conclusion

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

- RExcel allows for instantaneous transfer of pivot tables from Excel to R
- once the table is transferred the user has the full power of R to do statistical analysis
- R packages such as `vcd` provide useful graphics that are not available in Excel

# References and Links

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

Baier, T. and Neuwirth, E. (2007), Excel :: COM :: R.  
*Computational Statistics* 22/1 91-108

Dawson, Robert J. MacG. (1995) The "Unusual Episode" Data Revisited.  
*Journal of Statistics Education* v.3, n.3

Heiberger, R. and Neuwirth, E. (2009), *R Through Excel*. Springer Verlag  
Hornik, K. The Strucplot Framework: Visualizing Multi-way Contingency Tables  
with vcd. *Journal of Statistical Software, American Statistical Association*  
vol. 17(i03).

Jelen B. and Alexander, M. (2005), *Pivot Table Data Crunching*. Que Publishing  
Report on the Loss of the 'Titanic' (S.S.) (1990),  
*British Board of Trade Inquiry Report* (reprint). Allan Sutton Publishing

RExcelInstaller at CRAN

<http://cran.r-project.org/web/packages/RExcelInstaller/index.html>

RExcel Homepage

<http://rcom.univie.ac.at/>

Titanic Dataset

<http://www.amstat.org/publications/jse/v3n3/datasets.dawson.html>

# R Code for Creating Mosaic Plots

Export Pivot  
Table to R  
Using RExcel

Halbert,  
Heiberger,  
Neuwirth

About RExcel

About Pivot  
Tables

Why Transfer  
to R?

How it Works

Example

Conclusion

Appendix

```
mosaic(~ Class + Survive,  
       direction = c("v", "h"),  
       data = test,  
       highlighting = "Survive",  
       highlighting_fill = c("salmon", "skyblue")  
)  
mosaic(~ Class + Survive + Sex + Age,  
       direction = c("v", "h"),  
       data = test,  
       highlighting = "Survive",  
       highlighting_fill = c("salmon", "skyblue")  
)
```