

Feature Tour

By Benedict Escoto



This paper is produced mechanically as part of FAViR.
See <http://www.favir.net> for more information.

Abstract

This paper intended as an initial tour of the functions included in the FAViR package. It itself is an example of a FAViR paper that uses these functions.

1 Introduction

The R package component of FAViR includes several formatting functions which make writing FAViR papers more convenient. They are covered in the package manual pages but some only make sense inside a FAViR paper. This paper demonstrates each of these functions and can be used as a quick-reference or a feature tour of the FAViR package.

When reading the PDF version of this paper, please follow along in the associated Sweave (.Rnw) version.

2 LaTeX

LaTeX commands are available inside the LaTeX (non-R) sections of a FAViR paper. In particular:

- *text in italics*
- **boldface text**
- `typewriter font`
- tiny script footnote small normal large Large LARGE huge Huge

This text is centered.

1. This list is numbered.

2. And it has equations like $3 \leq 4$ and $1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} + \cdots = \sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}$

You can also have numbered, centered equations, like equation (1) below.

$$K(x) = \begin{cases} \frac{3}{4}(1 - x^2/5)/\sqrt{5} & |x| < \sqrt{5} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

LaTeX has far too many features to demonstrate here, so please consult the internet for general LaTeX help.

3 Sweave

Sweave allows R code to be included in a paper and automatically executed. Below is the most basic Sweave mode, where both R commands and their output are printed:

```
> 2 + 2
```

```
[1] 4
```

```
> 1:10
```

```
[1] 1 2 3 4 5 6 7 8 9 10
```

However, the output of this mode generally doesn't look very polished. A more powerful technique is not to display the R commands, but have them generate LaTeX output directly. This is how most code segments in FAViR papers are run.

This LaTeX is generated by R

$$\sum_{i=1}^{100} \log(i) \approx 363.739375555563$$

Note the double backslashes in the above R code. When that code is evaluated by R, each double backslash becomes a single backslash which is then evaluated by LaTeX.

It's also possible to run R without displaying the results or even in the *middle* of a sentence.

4 FAViR Data Frames

This section is about presenting tabular data in R data frames. First we'll define a simple data frame and then progressively refine its presentation by using FAViR functions.

An easy way to present a data frame is by using basic Sweave mode:

```
> df <- data.frame(year = 2000:2008, prem.a = 2e+05 * rnorm(9,
+   1, 0.1), loss.a = 130000 * rnorm(9, 1, 0.2), loss.ratio.a = 0,
+   prem.b = 50000 * rnorm(9, 1, 0.1), loss.b = 30000 * rnorm(9,
+   1, 0.2), loss.ratio.b = 0)
> df$loss.ratio.a <- df$loss.a/df$prem.a
> df$loss.ratio.b <- df$loss.b/df$prem.b
> df
```

	year	prem.a	loss.a	loss.ratio.a	prem.b	loss.b	loss.ratio.b
1	2000	203829.4	140564.22	0.6896172	44599.00	30699.09	0.6883358
2	2001	215568.4	126918.64	0.5887629	48172.25	33772.76	0.7010832
3	2002	179312.0	170276.48	0.9496099	45655.11	32995.04	0.7227019
4	2003	182152.7	146779.78	0.8058062	49522.29	23169.18	0.4678536
5	2004	187813.9	97623.26	0.5197873	47443.78	24242.09	0.5109644
6	2005	200782.6	149990.38	0.7470289	50861.83	27052.41	0.5318805
7	2006	193008.6	102532.45	0.5312324	45875.75	30432.12	0.6633597
8	2007	243428.6	188986.33	0.7763523	52151.08	35259.23	0.6760977
9	2008	216966.7	162186.45	0.7475177	60323.68	25991.45	0.4308665

However the results are obviously rough. We can improve on this by using a FavirDF object:

year	prem.a	loss.a	loss.ratio.a	prem.b	loss.b	loss.ratio.b
2,000	203,829	140,564	1	44,599	30,699	1
2,001	215,568	126,919	1	48,172	33,773	1
2,002	179,312	170,276	1	45,655	32,995	1
2,003	182,153	146,780	1	49,522	23,169	0
2,004	187,814	97,623	1	47,444	24,242	1
2,005	200,783	149,990	1	50,862	27,052	1
2,006	193,009	102,532	1	45,876	30,432	1
2,007	243,429	188,986	1	52,151	35,259	1
2,008	216,967	162,186	1	60,324	25,991	0

Figure 1:

The results are now displayed in a figure. LaTeX does not necessarily display figures where they are defined—it uses an algorithm to place them where it thinks they will fit. It's good practice to add labels and captions to each FavirDF, so it's clear which table is being referred to.

year	prem.a	loss.a	loss.ratio.a	prem.b	loss.b	loss.ratio.b
2,000	203,829	140,564	1	44,599	30,699	1
2,001	215,568	126,919	1	48,172	33,773	1
2,002	179,312	170,276	1	45,655	32,995	1
2,003	182,153	146,780	1	49,522	23,169	0
2,004	187,814	97,623	1	47,444	24,242	1
2,005	200,783	149,990	1	50,862	27,052	1
2,006	193,009	102,532	1	45,876	30,432	1
2,007	243,429	188,986	1	52,151	35,259	1
2,008	216,967	162,186	1	60,324	25,991	0

Figure 2: This FavirDF has a label and caption.

Figure 2 has a label and caption. It arguably looks better than the simple Sweave version, but the headers are chosen for R programming convenience and are not ideal for presentation.

4.1 Table Headings

We can change the headings by using the `FieldHeadings` function. Figure 3 has better column headings.

In figure 4 we do even better by grouping the lines of business.

Calendar Year	Earned Premium Line A	Ultimate Loss Line A	Loss Ratio Line A	Earned Premium Line B	Ultimate Loss Line B	Loss Ratio Line B
2,000	203,829	140,564	1	44,599	30,699	1
2,001	215,568	126,919	1	48,172	33,773	1
2,002	179,312	170,276	1	45,655	32,995	1
2,003	182,153	146,780	1	49,522	23,169	0
2,004	187,814	97,623	1	47,444	24,242	1
2,005	200,783	149,990	1	50,862	27,052	1
2,006	193,009	102,532	1	45,876	30,432	1
2,007	243,429	188,986	1	52,151	35,259	1
2,008	216,967	162,186	1	60,324	25,991	0

Figure 3: Table with better column headings

Calendar Year	Line of Business A			Line of Business B		
	Earned Premium	Ultimate Loss	Loss Ratio	Earned Premium	Ultimate Loss	Loss Ratio
2,000	203,829	140,564	1	44,599	30,699	1
2,001	215,568	126,919	1	48,172	33,773	1
2,002	179,312	170,276	1	45,655	32,995	1
2,003	182,153	146,780	1	49,522	23,169	0
2,004	187,814	97,623	1	47,444	24,242	1
2,005	200,783	149,990	1	50,862	27,052	1
2,006	193,009	102,532	1	45,876	30,432	1
2,007	243,429	188,986	1	52,151	35,259	1
2,008	216,967	162,186	1	60,324	25,991	0

Figure 4: Table with group headings

4.2 Entry Formatting

This table is starting to look decent, but the numbers are not formatted correctly. You can specify arbitrary formatting functions per field, but several standard ones are built into the FAViR package. In figure 5 the loss ratio and years columns have been reformatted.

Calendar Year	Line of Business A			Line of Business B		
	Earned Premium	Ultimate Loss	Loss Ratio (%)	Earned Premium	Ultimate Loss	Loss Ratio (%)
2000	203,829	140,564	69.0	44,599	30,699	68.8
2001	215,568	126,919	58.9	48,172	33,773	70.1
2002	179,312	170,276	95.0	45,655	32,995	72.3
2003	182,153	146,780	80.6	49,522	23,169	46.8
2004	187,814	97,623	52.0	47,444	24,242	51.1
2005	200,783	149,990	74.7	50,862	27,052	53.2
2006	193,009	102,532	53.1	45,876	30,432	66.3
2007	243,429	188,986	77.6	52,151	35,259	67.6
2008	216,967	162,186	74.8	60,324	25,991	43.1

Figure 5: Table with formatted entries

Note also that LaTeX moved some of the earlier figures onto other pages because there were “too many” figures and not enough text. At the start of this section we have manually forced all figures to be placed in order to start fresh.

4.3 Final Table

The last change we will make is to add a summary row. Figure 6 demonstrates this.

Calendar Year	Line of Business A			Line of Business B		
	Earned Premium	Ultimate Loss	Loss Ratio (%)	Earned Premium	Ultimate Loss	Loss Ratio (%)
2000	203,829	140,564	69.0	44,599	30,699	68.8
2001	215,568	126,919	58.9	48,172	33,773	70.1
2002	179,312	170,276	95.0	45,655	32,995	72.3
2003	182,153	146,780	80.6	49,522	23,169	46.8
2004	187,814	97,623	52.0	47,444	24,242	51.1
2005	200,783	149,990	74.7	50,862	27,052	53.2
2006	193,009	102,532	53.1	45,876	30,432	66.3
2007	243,429	188,986	77.6	52,151	35,259	67.6
2008	216,967	162,186	74.8	60,324	25,991	43.1
Total	1,822,863	1,285,858	70.5	444,605	263,613	59.3

Figure 6: Final table with summary row

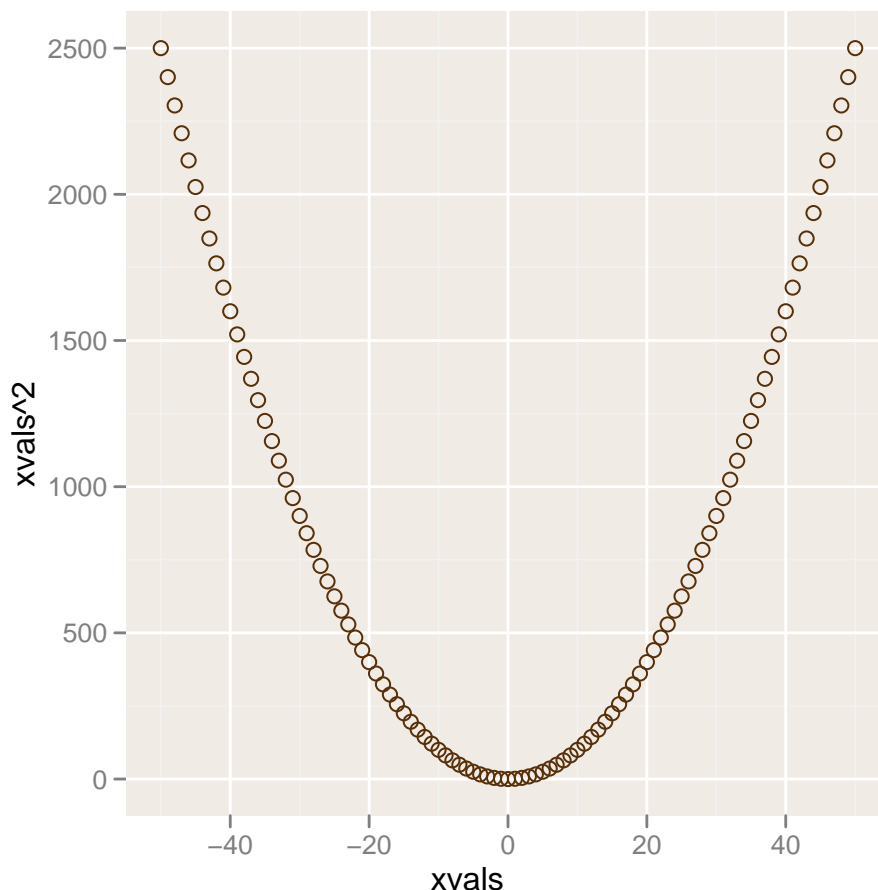


Figure 7: Basic graph with ggplot2

5 Graphics

ggplot2 is an advanced graphing package for R. The “gg” stands for “grammar of graphics”. The idea is that you just need to learn a few basic graphing concepts (such as geoms, aesthetics, and scales) and then you can combine them in myriad ways to create great custom plots. There is lots of documentation for *ggplot2*; this paper just demonstrates the FAViR functions which include them into a paper.

Figure 7 is a simple graph made with *ggplot2*.

The size can be adjusted with the **height** and **width** parameters as shown in figure 8.

Figure 9 shows how you can easily put multiple graphs into different viewports with the **IncludeGrid** function. However, if the graphs all have comparable values in them, it’s probably better to use *ggplot2*’s powerful faceting functionality.

Our final feature demonstration is to automatically include some legal boilerplate in the

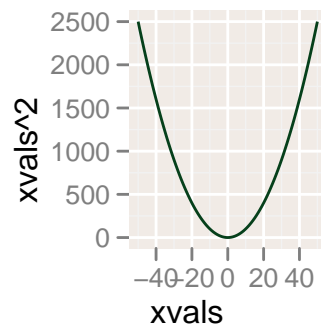


Figure 8: Smaller Graph

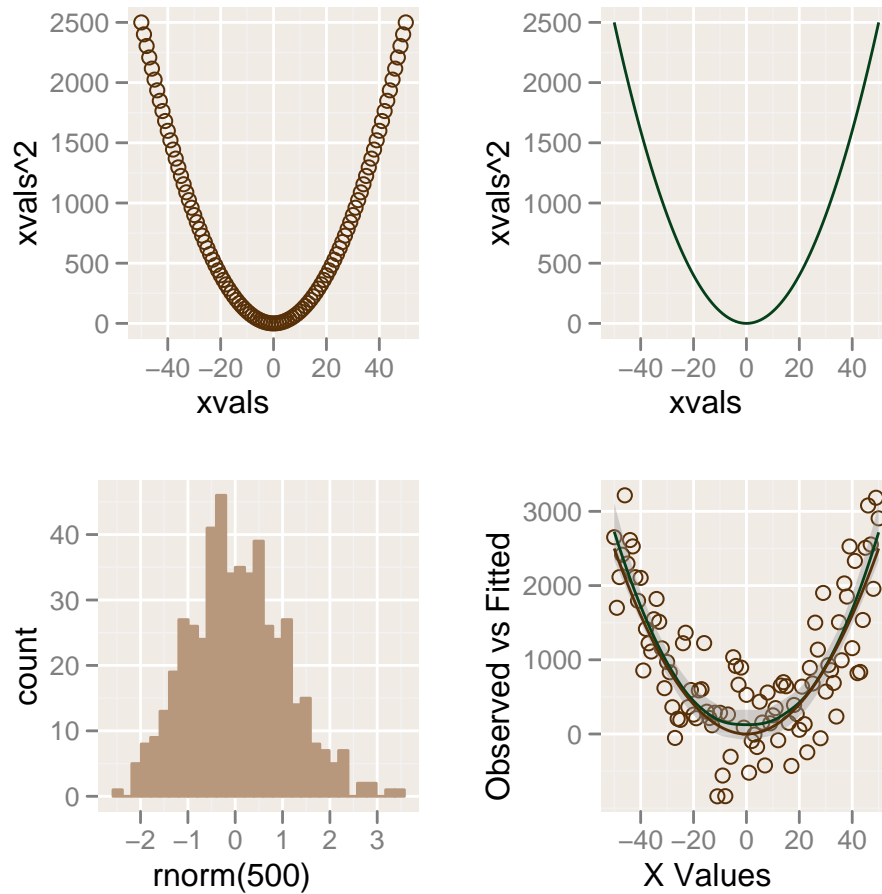


Figure 9: Four separate plots in one grid

next section.

6 Legal

Copyright © 2010 FAViR Project

This paper is part of the FAViR project. It and the R source code used to produce it are freely distributable under the GNU General Public License. See <http://www.favir.net> for more information on FAViR or to download the source code for this paper.

No warranty, express or implied, is provided with this paper. This paper is intended for educational purposes only and should not be used to violate anti-trust law. The authors and FAViR editors do not necessarily endorse the information or techniques in this paper and assume no responsibility for their accuracy.