



R news and tutorials contributed by (573) R bloggers

- [Home](#)
- [About](#)
- [RSS](#)
- [add your blog!](#)
- [R jobs](#)
- [Contact us](#)

Welcome!


Follow [@rbloggers](#) 21.3K

Here you will find daily **news and tutorials about R**, contributed by over 573 bloggers. There are many ways to **follow us** - [By e-mail:](#)

18324 readers


BY FEEDBURNER

[On Facebook:](#)



R blogg...
25k likes

Be the first of your friends to like this



If you are an **R blogger yourself** you are invited to [add your own R content feed to this site](#) (Non-English R bloggers should add themselves- [here](#))

[Jobs for R-users](#)

- [Senior Analyst for Health Insurance Intelligence with R & Shiny \(@Zurich\)](#)
- [Scientist / Sr. Scientist](#)
- [Sr. Manager of Content Analytics at Amazon.com \(@Seattle\)](#)
- [Data Scientist @Burlington](#)
- [Associate Director of Admissions Strategy and Analysis \(@Maine\)](#)

Search & Hit Enter

Popular

Searches

- [heatmap](#)
- [web scraping](#)
- [maps](#)
- [hadoop](#)
- [shiny](#)
- [twitterR](#)
- [ggplot2](#)
- [boxplot](#)
- [trading](#)
- [animation](#)
- [excel](#)
- [Time series](#)
- [finance](#)
- [ggplot](#)
- [latex](#)
- [pca](#)
- [googlevis](#)
- [eclipse](#)
- [quantmod](#)
- [rstudio](#)
- [market research](#)
- [tutorial](#)
- [how to import image file to R](#)
- [remdr](#)
- [knitr](#)
- [rattle](#)
- [gis](#)
- [coplot](#)
- [sweave](#)
- [ecdf](#)

Recent Posts

- [Deaths in the Netherlands by cause and age](#)
- [The Workflow of Infinite Shame, and other stories from the R Summit](#)
- [What is a good Sharpe ratio?](#)
- [Open Analytics @ UseR! 2015](#)
- [When did same-sex marriage laws become effective by state?](#)
- [Bio7 2.2 for Windows, Linux and MacOSX released](#)
- [Rperform: Obtaining quantitative metrics across R package versions](#)
- [A bit about Win-Vector LLC](#)
- [R at Microsoft](#)
- [Introduction to Monte Carlo methods with R and Bayesian Essentials with R](#)
- [An Attempt to Understand Boosting Algorithm\(s\)](#)
- [remote update, paper & poster](#)
- [Geomorph update 2.1.6 Now Available!](#)
- [KDD Cup 2015: The story of how I built hundreds of predictive](#)

[models....And got so close, yet so far away from 1st place!](#)

- [Working with the RStudio CRAN logs](#)

Other sites

- [Jobs for R-users](#)
- [Statistics of Israel](#)
- [SAS blogs](#)

A nifty area plot (or a bootleg of a ggplot geom)

September 17, 2013

By [beckmw](#)

Tweet

(This article was first published on [R is my friend » R](#), and kindly contributed to [R-bloggers](#))

The ideas for most of my blogs usually come from half-baked attempts to create some neat or useful feature that hasn't been implemented in R. These ideas might come from some analysis I've used in my own research or from some other creation meant to save time. More often than not, my blogs are motivated by data visualization techniques meant to provide useful ways of thinking about analysis results or comparing pieces of information. I was particularly excited (anxious?) after I came across this really [neat graphic](#) that ambitiously portrays the progression of civilization since ca. BC 2000. The original image was created John B. Sparks and was first printed by Rand McNally in 1931.¹ The 'Histomap' illustrates the rise and fall of various empires and civilizations through an increasing time series up to present day. Color-coded chunks at each time step indicate the relative importance or dominance of each group. Although the methodology by which 'relative dominance' was determined is unclear, the map provides an impressive synopsis of human history.



Fig: Snippet of the histomap. See the footnote for full link.¹

Historical significance aside, I couldn't help but wonder if a similar figure could be reproduced in R (I am not a historian). I started work on a function to graphically display the relative importance of categorical data across a finite time series. After a few hours of working, I soon realized that plot functions in the [ggplot2](#) package can already accomplish this task. Specifically, the [geom_area](#) 'geom' provides a 'continuous analog of a stacked bar chart, and can be used to show how composition of the whole varies over the range of x'.² I wasn't the least bit surprised that this functionality was already available in [ggplot2](#). Rather than scrapping my function entirely, I decided to stay the course in the naive hope that a stand-alone function that was independent of [ggplot2](#) might be useful. Consider this blog my attempt at ripping-off some ideas from Hadley Wickham.

The first question that came to mind when I started the function was the type of data to illustrate. The data should illustrate changes in relative values (abundance?) for different categories or groups across time. The only assumption is that the relative values are all positive. Here's how I created some sample data:

```
1 #create data
2 set.seed(3)
3
4 #time steps
5 t.step<-seq(0,20)
6
7 #group names
8 grps<-letters[1:10]
9
10 #random data for group values across time
11 grp.dat<-runif(length(t.step)*length(grps),5,
12
13 #create data frame for use with plot
14 grp.dat<-matrix(grp.dat,nrow=length(t.step),n
```

```
15 | grp.dat<-data.frame(grp.dat,row.names=t.step)
16 | names(grp.dat)<-grps
```

The code creates random data from a uniform distribution for ten groups of variables across twenty time steps. The approach is similar to the method used in my blog about a [nifty line plot](#). The data defined here can also be used with the line plot as an alternative approach for visualization. The only difference between this data format and the latter is that the time steps in this example are the row names, rather than a separate column. The difference is trivial but in hindsight I should have kept them the same for compatibility between plotting functions. The data have the following form for the first four groups and first three time steps.

	a	b	c	d
0	6.7	5.2	6.7	6.0
1	13.1	6.3	10.7	12.7
2	8.8	5.9	9.2	8.0
3	8.3	7.4	7.7	12.7

The plotting function, named `plot.area`, can be imported and implemented as follows (or just copy from the link):

```
1 | source("https://gist.github.com/fawda123/65895")
2 |
3 | plot.area(grp.dat)
```

Fig: The `plot.area` function in action.

The function indicates the relative values of each group as a proportion from 0-1 by default. The function arguments are as follows:

x	data frame with input data, row names are time steps
col	character string of at least one color vector for the groups, passed to <code>colorRampPalette</code> , default is 'lightblue' to 'green'
horiz	logical indicating if time steps are arranged horizontally, default F
prop	logical indicating if group values are displayed as relative proportions from 0–1, default T
stp.ln	logical indicating if parallel lines are plotted indicating the time steps, default T
grp.ln	logical indicating if lines are plotted connecting values of individual groups, default T
axs.cex	font size for axis values, default 1
axs.lab	logical indicating if labels are plotted on each axis, default T
lab.cex	font size for axis labels, default 1
names	character string of length three indicating axis names, default 'Group', 'Step', 'Value'
...	additional arguments to be passed to <code>par</code>

I arbitrarily chose the color scheme as a ramp from light blue to green. Any combination of color values as input to `colorRampPalette` can be used. Individual colors for each group will be used if the number of input colors is equal to the number of groups. Here's an example of another color ramp.

```
1 | plot.area(grp.dat,col=c('red','lightgreen','pu
```

Fig: The `plot.area` function using a color ramp from red, light green, to purple.

The function also allows customization of the lines that connect time steps and groups.

```
1 | plot.area(grp.dat,col=c('red','lightgreen','pu
2 |
3 | plot.area(grp.dat,col=c('red','lightgreen','pu
```

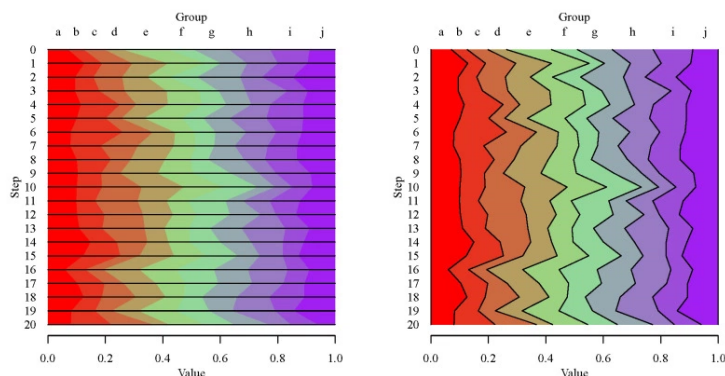


Fig: The `plot.area` function with different line customizations.

Finally, the `prop` argument can be used to retain the original values of the data and the `horiz` argument can be used to plot groups horizontally.

```
1 | plot.area(grp.dat,prop=F)
2 |
3 | plot.area(grp.dat,prop=F,horiz=T)
```

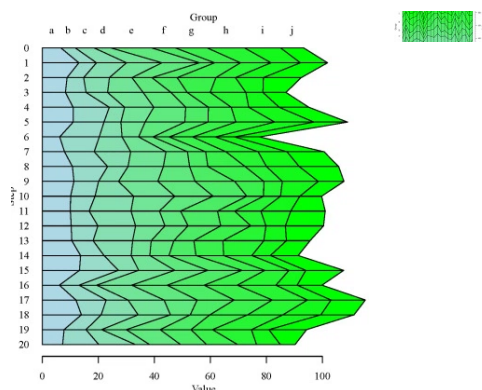


Fig: The `plot.area` function with changing arguments for `prop` and `horiz`.

Now is a useful time to illustrate how these graphs can be replicated in `ggplot2`. After some quick reshaping of our data, we can create a similar plot as above (full code [here](#)):

```
1 | require(ggplot2)
2 | require(reshape)
3 |
4 | #reshape the data
5 | p.dat<-data.frame(step=row.names(grp.dat),grp
6 | p.dat<-melt(p.dat,id='step')
7 | p.dat$step<-as.numeric(p.dat$step)
8 |
9 | #create plots
10 | p<-ggplot(p.dat,aes(x=step,y=value)) + theme(
11 | p + geom_area(aes(fill=variable))
12 | p + geom_area(aes(fill=variable),position='fi
```

Fig: The `ggplot2` approach to plotting the data.

Same plots, right? My function is practically useless given the tools in `ggplot2`. However, the `plot.area` function is completely independent, allowing for easy manipulation of the source code. I'll leave it up to you to decide which approach is most useful.

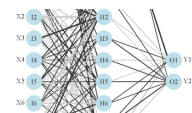
-Marcus

¹http://www.slate.com/blogs/the_vault/2013/08/12/the_1931_histomap_the_entire_history_of_the_world_distilled_into_a_single.html

²http://docs.ggplot2.org/current/geom_area.html



Related



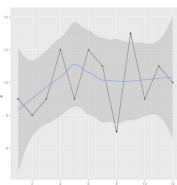
Sensitivity analysis for neural networks

In "R bloggers"

R Tips for Quantitative Trading

The article below is an updated version of an article I wrote for R-Bloggers in August 2010. As a first post I

In "R bloggers"



Better than Average

Similar post

Tweet

To leave a comment for the author, please follow the link and comment on his blog: [R is my friend » R](#).

R-bloggers.com offers [daily e-mail updates](#) about [R](#) news and [tutorials](#) on topics such as: visualization ([ggplot2](#), [Boxplots](#), [maps](#), [animation](#)), programming ([RStudio](#), [Sweave](#), [LaTeX](#), [SQL](#), [Eclipse](#), [git](#), [hadoop](#), [Web Scraping](#)) statistics ([regression](#), [PCA](#), [time series](#), [trading](#)) and more...

If you got this far, why not [subscribe for updates](#) from the site? Choose your flavor: [e-mail](#), [twitter](#), [RSS](#), or [facebook](#)...

Tweet

Comments are closed.

Top 3 Posts from the past 2 days

- [R at Microsoft](#)
- [In-depth introduction to machine learning in 15 hours of expert videos](#)
- [The Single Most Important Skill for a Data Scientist](#)

Search & Hit Enter

Top 9 articles of the week


1. [Scatterplots](#)
2. [In-depth introduction to machine learning in 15 hours of expert videos](#)
3. [Installing R packages](#)
4. [The Single Most Important Skill for a Data Scientist](#)
5. [Illustrated Guide to ROC and AUC](#)
6. [Using apply, sapply, lapply in R](#)
7. [Network analysis with igraph](#)
8. [R vs Python: Survival Analysis with Plotly](#)
9. [KDD Cup 2015: The story of how I built hundreds of predictive models....And got so close, yet so far away from 1st place!](#)

Sponsors



Highland Statistics Ltd

Zero Inflated Models & GLMM
 Beginner's Guide to GAM
 Beginner's Guide to GLM & GLMM
 Beginner's Guide to GAMM



R training
 R consulting

QUANTIDE
 knowledge from data

R Studio
 open source & enterprise ready
 professional software for R

REVOLUTION
 ANALYTICS

R for the Enterprise

www.revolutionanalytics.com

Werden Sie zum Expe[R]ten mit der
 R-Akademie von

eoda
 daten • wissen • nutzen

Beratung | Software
 Training | Lösungen

plotly online R graphing

[Plotly: collaborative, publication-quality graphing.](#)

STATISTICS
 VIEWS

Bringing Statistics Together

NYC DATA SCIENCE ACADEMY
 Jun 1st - Aug 21st, 2015 | Full-Time Program

12-WEEK DATA SCIENCE BOOTCAMP
 Hands-on R, Python & Hadoop Training

Mentorship from Top Data Scientists
 Job Placement at Our 300+ Hiring Firms
 Scholarships & Financial Aid Available

Deadline for Application May 1st // Apply Immediately



June 30 - July 3, 2015 Aalborg, Denmark

Save **25%** on
R Books

CRC Press
Taylor & Francis Group

Advanced R

John Fox

Use
Promo
Code
CZP40

Plus Free
Shipping

**Data Analysis
Modeling**

Cloud Solution

Free Sign Up!

dotplot

STATWORX

Consulting
Schulung
Data Mining

Mehr erfahren

**Try the FASTEST ML
for R**

Click for a Free Trial

**YOTTAMINE
ANALYTICS**

CLIMB THE BIG DATA BANDWAGON!

ALL-IN-ONE BIG DATA &
ANALYTICS SUITE

\$399

5-DAYS MONEY BACK GUARANTEE

simplilearn

ENROLL NOW

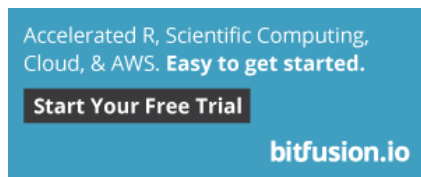
OPEN DATA SCIENCE CONFERENCE

Meet Data Science Thought Leaders
in Boston May 30th & 31st

Tickets Start at \$100 (early reg)

21 Workshops 76 Presentations #ODSC





Search & Hit Enter

[Full list of contributing R-bloggers](#)

[R-bloggers](#) was founded by [Tal Galili](#), with gratitude to the [R](#) community.

Is powered by [WordPress](#) using a [bavotasan.com](#) design.

Copyright © 2015 [R-bloggers](#). All Rights Reserved. [Terms and Conditions](#) for this website