

[Go](#)[Subscribe](#)

52793 readers

[Follow @rbloggers](#)

103K followers



R bloggers

[Like Page](#)

80K likes

Most viewed posts (weekly)

[PCA vs Autoencoders for Dimensionality Reduction](#)[5 Ways to Subset a Data Frame in R](#)[Which data science skills are important \(\\$50,000 increase in salary in 6-months\)](#)[Better Sentiment Analysis with sentiment.ai](#)[Self-documenting plots in ggplot2](#)[How to write the first for loop in R](#)

Graphically analyzing variable interactions in R

Posted on August 23, 2011 by Zach Mayer in [R bloggers](#) | 0 Comments


[This article was first published on [Modern Toolmaking](#), and kindly contributed to [R-bloggers](#)]. (You can report issue about the content on this page [here](#))

Want to share your content on R-bloggers? [click here](#) if you have a blog, or [here](#) if you don't.

[f](#) Share[t](#) Tweet

I studied Ecology as an undergraduate, which meant I spent a lot of time gathering and analyzing field data. One of the basic tools we used to look for relationships in a large set of variables was correlation and scatterplot matrices. Each of these requires a single line of code in R:

```
1 cor(iris[-5])
2 pairs(iris[-5], bg=iris$Species, pch=21)
```

0. Plain.R hosted with  by [GitHub](#)

[view raw](#)


1		Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
2	Sepal.Length	1.0000000	-0.1175698	0.8717538	0.8179411
3	Sepal.Width	-0.1175698	1.0000000	-0.4284401	-0.3661259
4	Petal.Length	0.8717538	-0.4284401	1.0000000	0.9628654
5	Petal.Width	0.8179411	-0.3661259	0.9628654	1.0000000


1. Plain result.R hosted with ❤️ by GitHub

view raw

How to Calculate a Cumulative Average in R

Sponsors

#earlconf



Full agenda is now live!

See the complete line up

datacamp

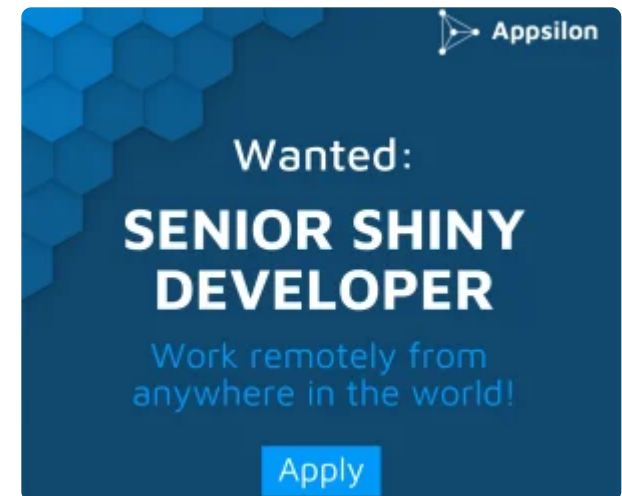
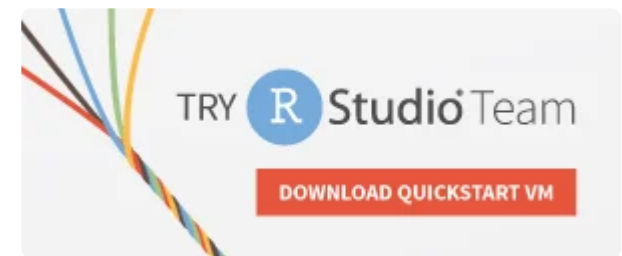
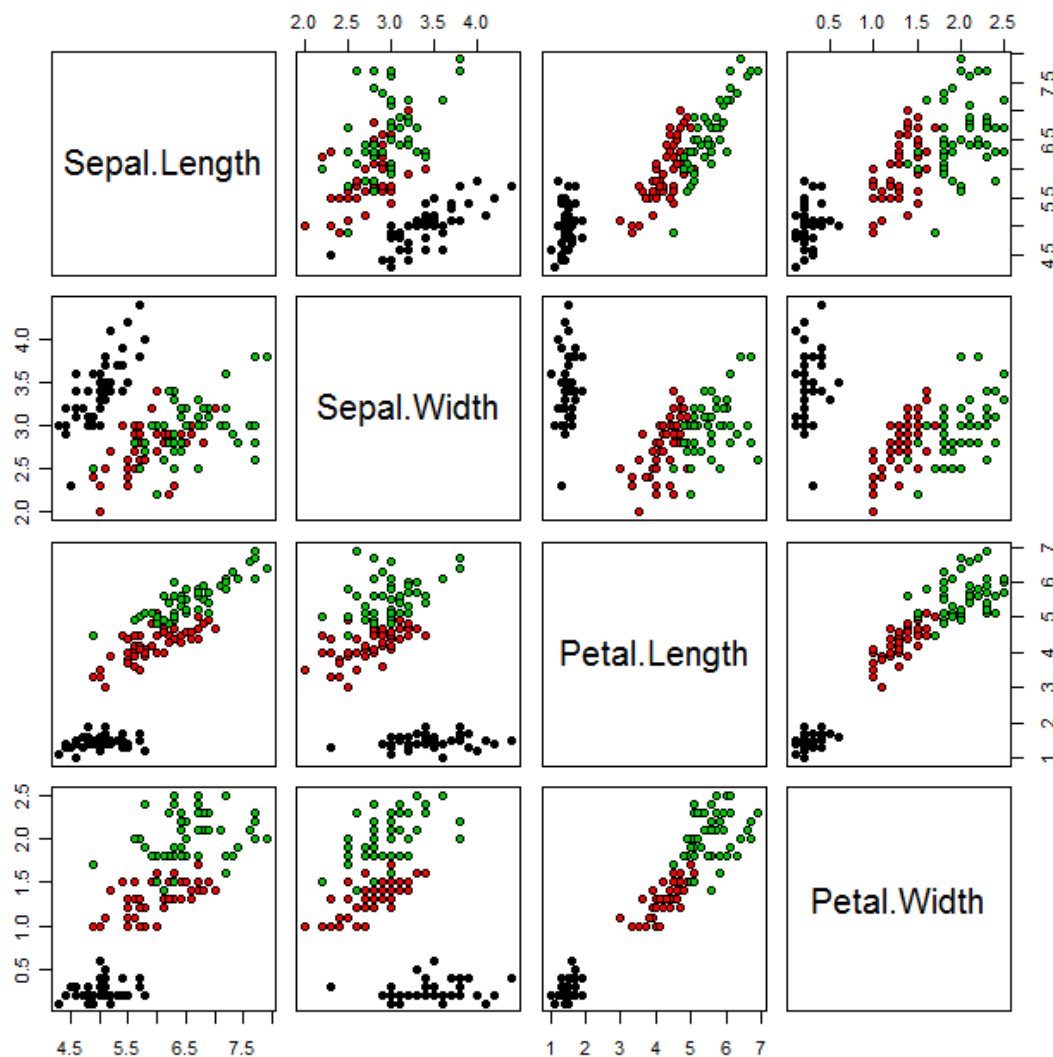
XP

Aug 15 - 31

LEARNER CHALLENGE

Win up to **\$1500*** by learning
data & earning points

Start Learning



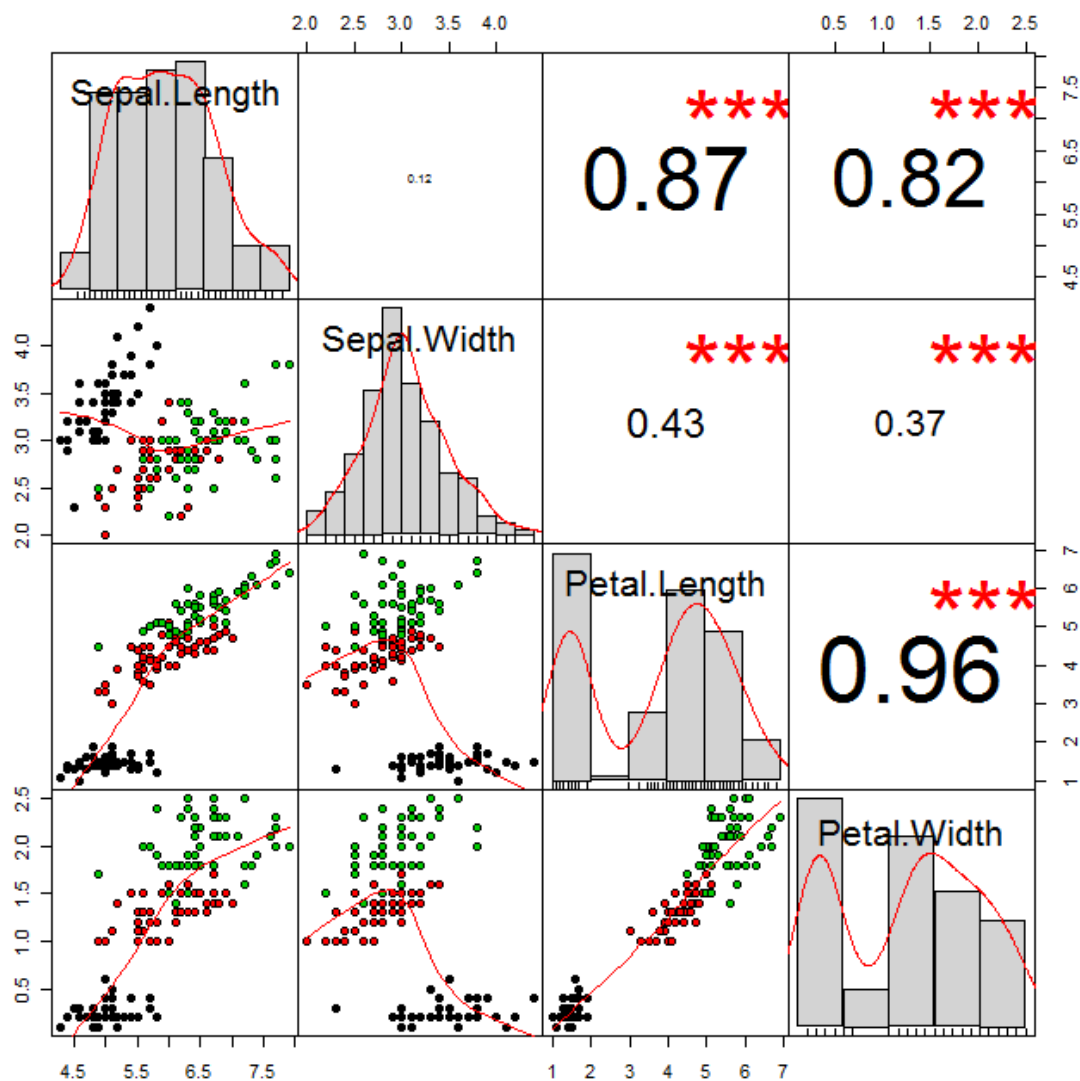
The 'pairs' function in R contains a lot of additional options, which can be used to make very informative plots. These options can get a little cumbersome, but fortunately several package authors have written wrapper functions that automatically enable some extra magic. Two such packages are [psych](#) and [PerformanceAnalytics](#). I happen to prefer the 1 liner from PerformanceAnalytics, but it's a matter of personal taste:

```
1 library(PerformanceAnalytics)
```

```
2 chart.Correlation(iris[-5], bg=iris$Species, pch=21)
```

2. Iris.R hosted with ❤ by GitHub

[view raw](#)



This chart contains a LOT of information: On the diagonal are the univariate distributions, plotted as histograms and kernel density plots. On the right of the

Managed RStudio Infrastructure



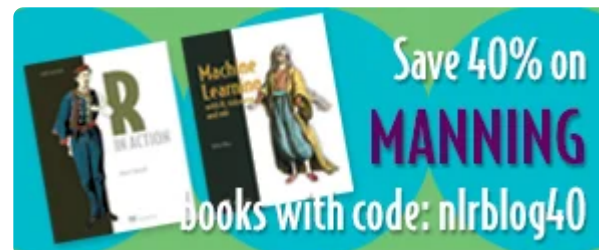
STATWORX

Data Science Service

Data Science | Consulting | Development | Training

SIX
SIGMA

SIX
SIGMA



diagonal are the pair-wise correlations, with red stars signifying significance levels. As the correlations get bigger the font size of the coefficient gets bigger. On the left side of the diagonal is the scatter-plot matrix, with loess smoothers in red to help illustrate the underlying relationship. This is one of my favorite plots in R, because it combines a large amount of information into one command and one easy to follow plot. In fact, this plot contains more information than is revealed by the 1st two commands in this post!

Of course, you can use this command on data from other domains besides Ecology. PerformanceAnalytics is intended for the analysis of financial data, so lets put it through its paces. First we download some financial data (a stock index, a bond index, and a gold index) from yahoo finance using quantmod, and then combine the daily close series of those indexes into one dataframe. I'm not 100% happy with the legend in the plot, but I wanted to show how the correlations between these indexes have changed over the years. I also skipped red (color #2) in the plots and in the legend, because the loess smoother is also red.

```
1 library(quantmod)
2 getSymbols(c('SPY', 'LQD', 'GLD'))
3 Data <- cbind(C1(SPY), C1(LQD), C1(GLD))
4 Year <- as.factor(format(index(Data), '%Y'))
5 chart.Correlation(Data, bg=seq(1:5)[Year], pch=21)
6 par(xpd=TRUE)
7 legend(0, 1, as.vector(unique(Year)), fill=seq(1:5))
```

3. Finance.R hosted with ❤ by GitHub

[view raw](#)

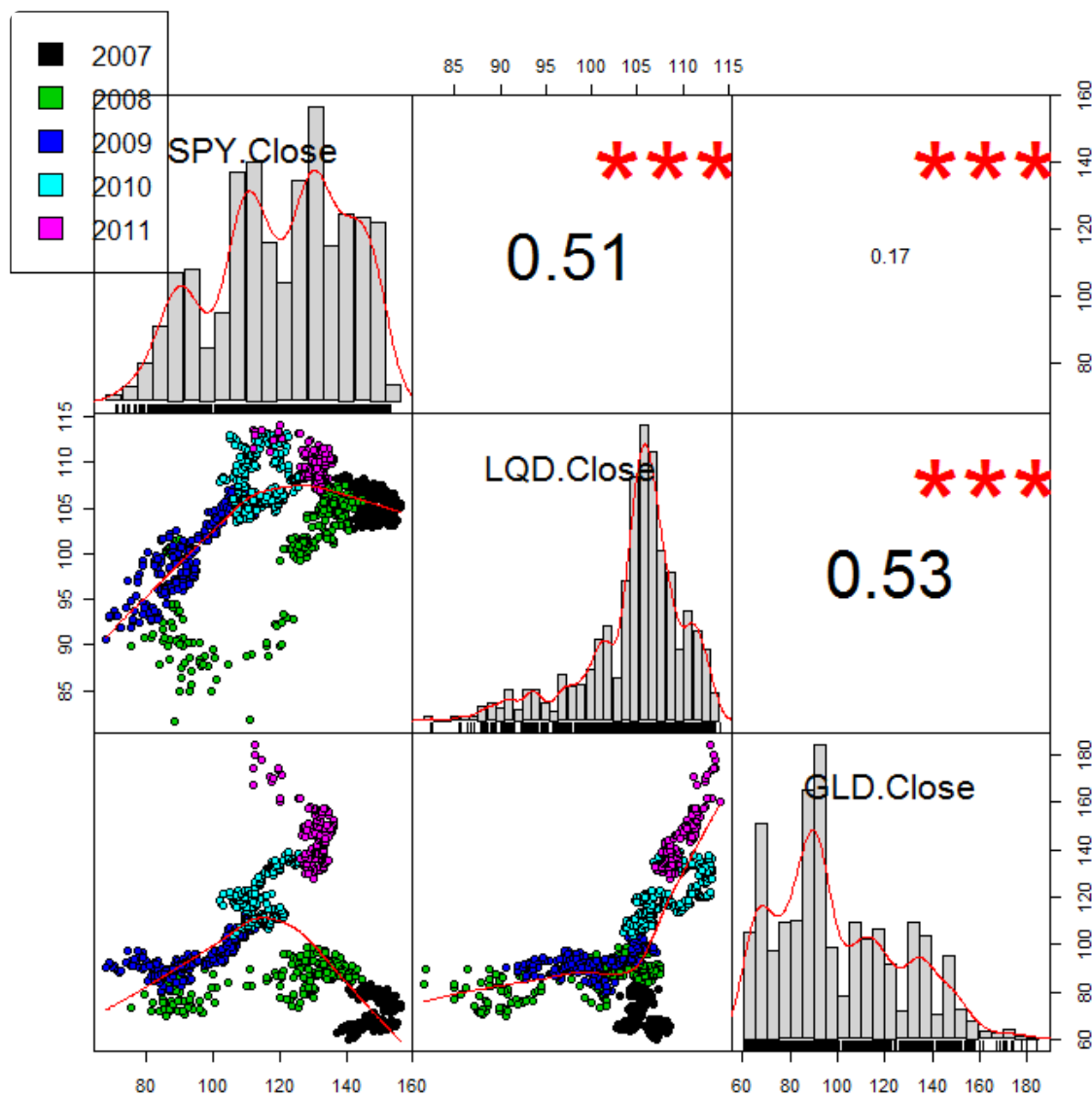


Our ads respect your privacy. Read our [Privacy Policy](#) page to learn more.

Contact us if you wish to help support R-bloggers, and place **your banner here**.

Recent Posts

Monte Carlo Analysis in R
Stock Market Predictions Next Week
{golem} 0.3.2 is now available
Convert column to categorical in R
Which data science skills are important
(\$50,000 increase in salary in 6-months)
A prerelease version of Jupyter Notebooks
and unleashing features in JupyterLab
Markov Switching Multifractal (MSM) model
using R package
Dashboard Framework Part 2: Running Shiny
in AWS Fargate with CDK
Something to note when using the merge
function in R
Better Sentiment Analysis with sentiment.ai
Self-documenting plots in ggplot2



Finally, I'd like to acknowledge [Stephen Turner](#) over on [cross-validated](#) for inspiring this post.

Data Challenges for R Users

simplevis: new & improved!

Checking the inputs of your R functions

Imputing missing values in R

Jobs for R-users

Junior Data Scientist / Quantitative economist

Senior Quantitative Analyst

R programmer

Data Scientist – CGIAR Excellence in

Agronomy (Ref No: DDG-

R4D/DS/1/CG/EA/06/20)

Data Analytics Auditor, Future of Audit Lead

@ London or Newcastle

python-bloggers.com (python/data-science news)

Object Oriented Programming in Python –
What and Why?

Dunn Index for K-Means Clustering Evaluation

Installing Python and Tensorflow with Jupyter

Notebook Configurations

How to Get Twitter Data using Python

Visualizations with Altair

Spelling Corrector Program in Python

Spelling Checker Program in Python

Full list of contributing R-bloggers

Related

More on Exploring Correlations in R

About a year ago I wrote a post about producing scatterplot matrices in R. These are handy for quickly getting a sense of August 28, 2012
Similar post

A Kaggle Dataset of R Package History for `rstudio::conf(2022)`

It's summer, and the long-awaited Rstudio conference for 2022 is only days away. Next week, a large number of R July 21, 2022
In "R bloggers"

Case Study: Modularizing a Package

Origins of deepdep
Separation of woodendesc
A new package is not always the answer
Modularization of tidysq August 3, 2022
In "R bloggers"

 Share

 Tweet


To **leave a comment** for the author, please follow the link and comment on their blog:
[Modern Toolmaking](#).

[R-bloggers.com](#) offers **daily e-mail updates** about R news and tutorials about [learning R](#) and many other topics. [Click here](#) if you're looking to post or find an R/data-science job.

Want to share your content on R-bloggers? [click here](#) if you have a blog, or [here](#) if you don't.

R Graphics

Archives

Select Month 

Other sites

[Jobs for R-users](#)

[SAS blogs](#)

[← Previous post](#)

[Next post →](#)

