

What is R?

PUBHLTH 597: Intro to Statistical Computing with R

Nicholas G Reich

UMass School of Public Health and Health Sciences


Division of Biostatistics and Epidemiology

people.umass.edu/nick

[@reichlab](mailto:nreich@reichlab.org)

R is a statistical software package

The R Project for Statistical Computing



About R
[What is R?](#)
[Contributors](#)
[Screenshots](#)
[What's new?](#)

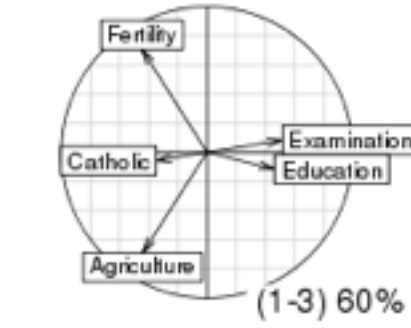
Download, Packages
[CRAN](#)

R Project
[Foundation](#)
[Members & Donors](#)
[Mailing Lists](#)
[Bug Tracking](#)
[Developer Page](#)
[Conferences](#)
[Search](#)

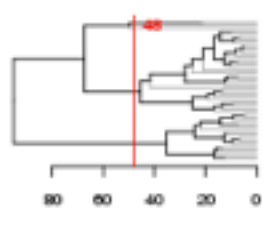
Documentation
[Manuals](#)
[FAQs](#)
[The R Journal](#)
[Wiki](#)
[Books](#)
[Certification](#)
[Other](#)

Misc
[Bioconductor](#)
[Related Projects](#)
[User Groups](#)
[Links](#)

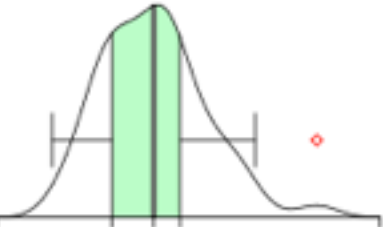
PCA 5 vars
`prcomp(x = data, cor = cor)`



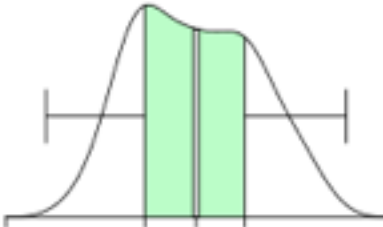
Clustering 4 groups



Factor 1 [41%]



Factor 3 [19%]



Getting Started:

- R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred [CRAN mirror](#).
- If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

News :

- **R version 2.15.1** (Roasted Marshmallows) has been released on 2012-06-22.
- **R version 2.14.2** (Gift-Getting Season) has been released on 2012-02-29.
- **The R Journal Vol.3/2** is available.
- **useR! 2012**, will take place at Vanderbilt University, Nashville Tennessee, USA, June 12-15, 2012.

R is a programming language

```
R version 2.15.1 (2012-06-22) -- "Roasted Marshmallows"  
Copyright (C) 2012 The R Foundation for Statistical Computing  
ISBN 3-900051-07-0  
Platform: x86_64-apple-darwin9.8.0/x86_64 (64-bit)
```

```
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.
```

```
  Natural language support but running in an English locale
```

```
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.
```

```
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.
```

```
> n <- 1000  
> means <- rep(NA, n)  
> for(i in 1:n){  
+   d <- rpois(10000, 2)  
+   means[i] = mean(d)  
+ }  
> hist(d)  
> hist(means)  
> summary(means)  
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   
 1.948  1.991   2.000   2.000   2.010   2.055   
> |
```

Three rules for R*

- Everything in R is an object.
- Everything that happens in R is a function.
- R is an interface.

* Adapted from John Chambers, useR!2014 keynote presentation

R is an open-source project

1 Package



sp

R is an open-source project

73 Packages

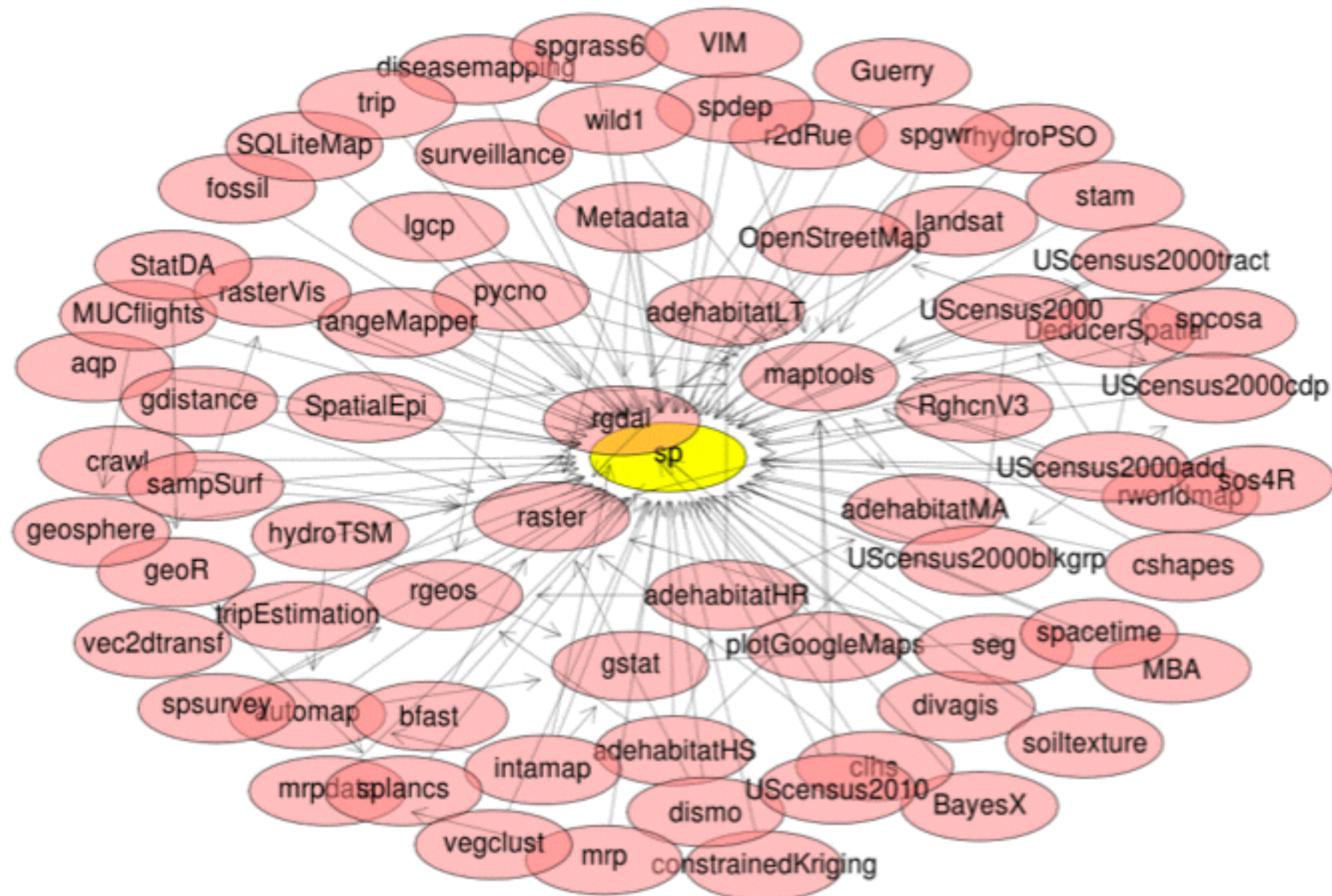
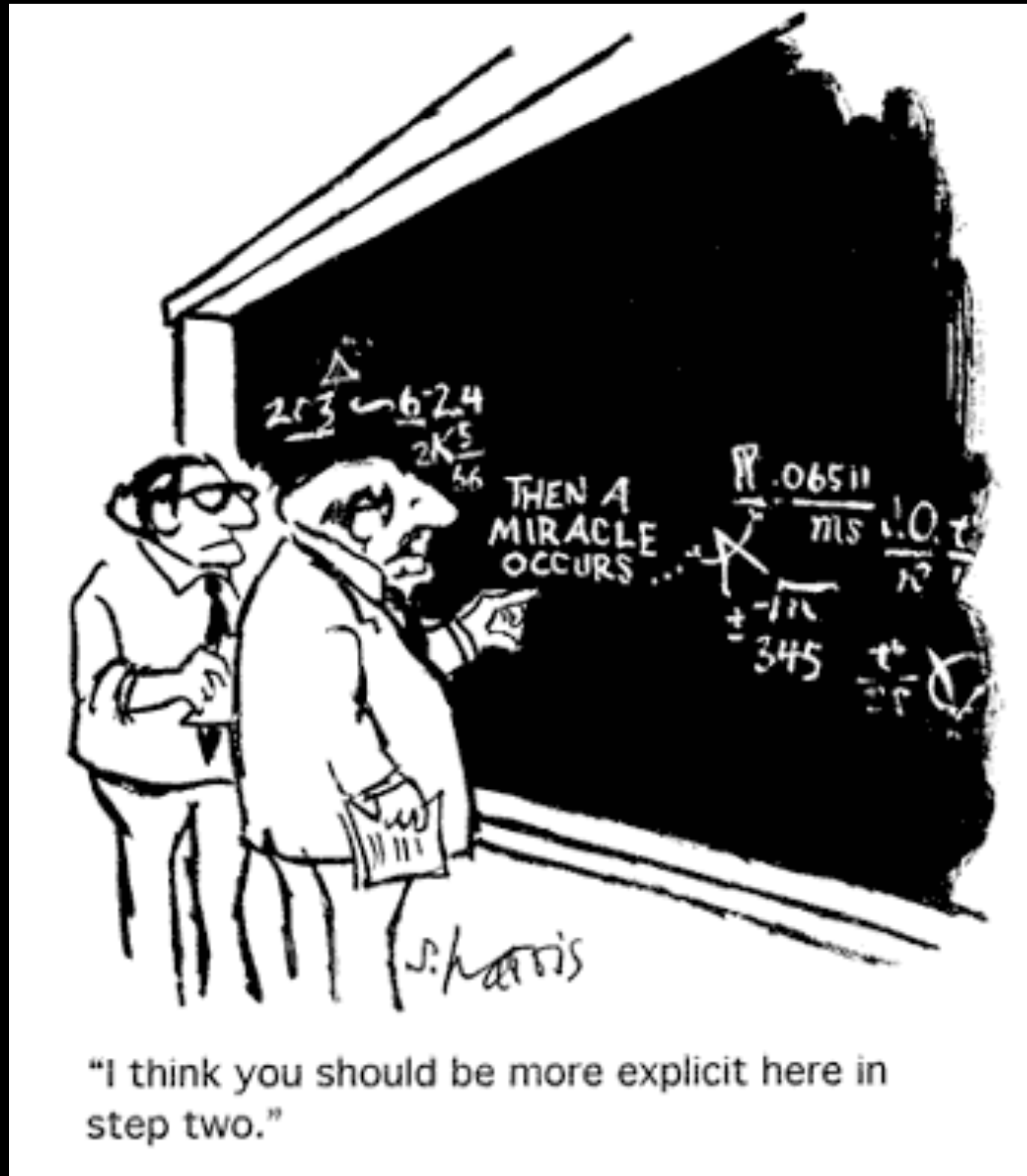


Figure acknowledgments to Barry Rowlingson's "Off Roading with Spatial" workshop at useR! 2012
<http://www.maths.lancs.ac.uk/~rowlings/Teaching/UseR2012/>
 The figures are available under a Creative Commons license.

116 Packages



R is a tool for reproducibility



R is hot

The New York Times

Business Computing

WORLD

U.S.

N.Y. / REGION

BUSINESS

TECHNOLOGY

SCIENCE

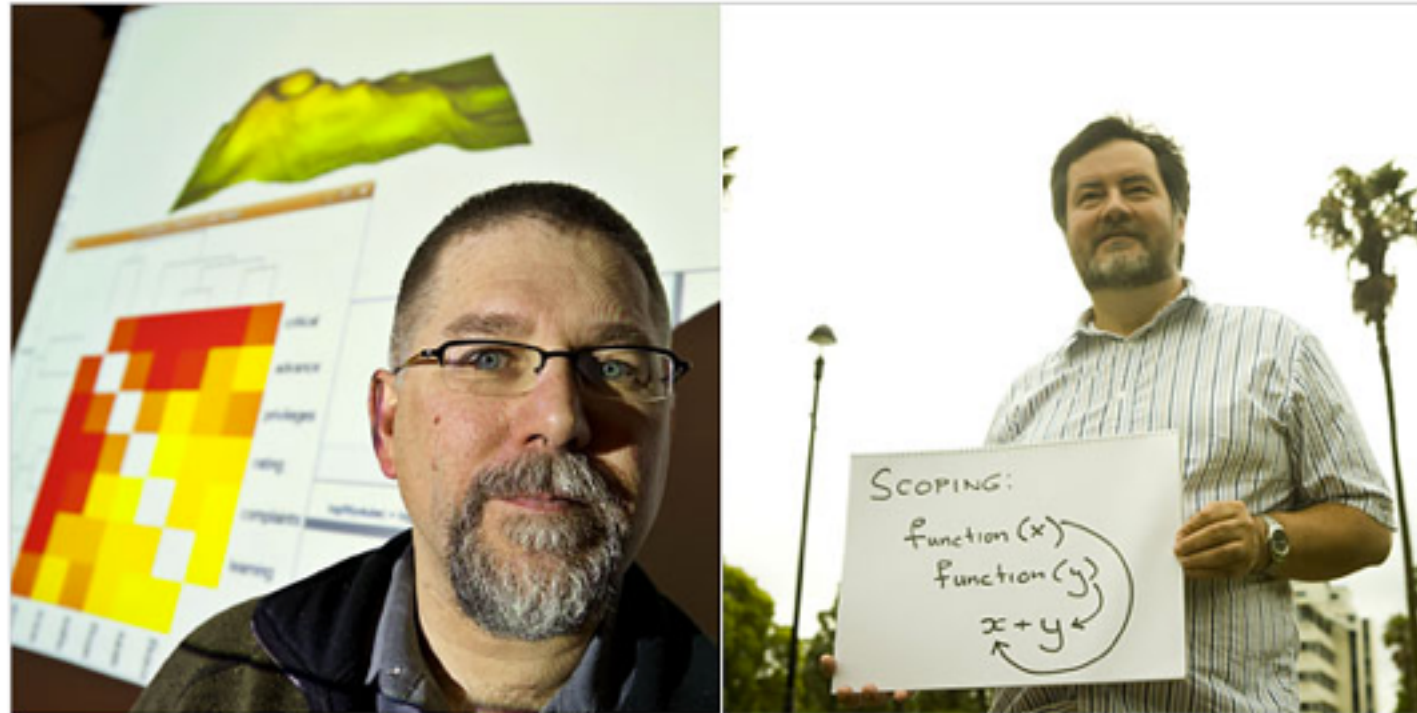
HEALTH

SPORTS

OPINION

ARTS

Data Analysts Captivated by R's Power



Stuart Isett for The New York Times

R first appeared in 1996, when the statistics professors Robert Gentleman, left, and Ross Ihaka released the code as a free software package.

By ASHLEE VANCE

Published: January 6, 2009

To some people R is just the 18th letter of the alphabet. To others, it's the rating on racy movies, a measure of an attic's insulation or what pirates in movies say.

Related

[Bits: R You Ready for R?](#)

[The R Project for Statistical Computing](#)

R is also the name of a popular programming language used by a growing number of data analysts inside corporations and academia. It is becoming their lingua franca partly because data mining has entered a

golden age, whether being used to set ad prices, find new drugs more quickly or fine-tune financial models. Companies as diverse as [Google](#), [Pfizer](#), [Merck](#), [Bank of America](#), the

TWITTER

LINKEDIN

SIGN IN TO E-MAIL

PRINT

REPRINTS

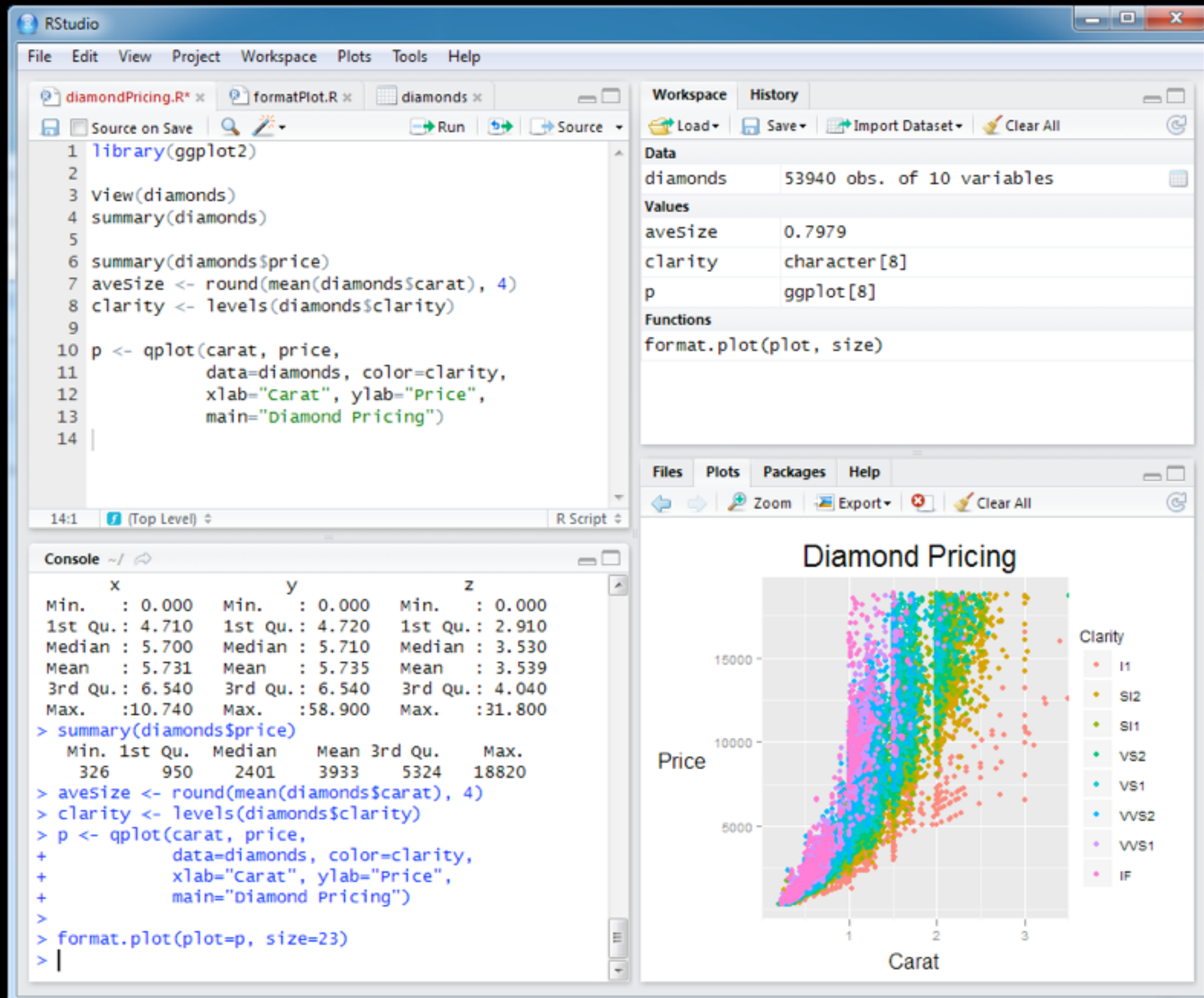
SHARE



R ...

- ... has a learning curve
- ... is best learned with others
- ... will occasionally give you a headache
- ... is worth the effort!

What is RStudio?



R resources

- google
- project homepage: r-project.org
- nerd list: stackoverflow.com/tags/r
- r-bloggers.com
- #rstats
- local: list.umass.edu/mailman/listinfo/r-users
- UMass student group: <http://tinyurl.com/oebsource>

Rules of the road

- Ask conceptual/content questions anytime.
- To keep things moving, if/when you get a coding error during an interactive session, please double-check for spelling and/or ask your neighbor how they got it to work.