

[Page One](#)
[Campus Computing News](#)
[Whither Windows? Windows Vista and Office 2007 - What is a poor student to do?](#)
[You're Invited: UNT Microsoft Vista/Office 2007 Launch Event](#)
[Windows Vista Learning Resources](#)
[UNT Learning Management System Evaluation Project Update](#)
[InHouse Series on Information Security](#)
[Voyage of the Enterprise](#)
[Today's Cartoon](#)
[RSS Matters](#)
[The Network Connection](#)
[Link of the Month](#)
[Helpdesk FYI](#)
[Short Courses](#)
[IRC News](#)
[Staff Activities](#)

Research and Statistical Support

University of North Texas

RSS Matters

You can link to the last RSS article here: [Statistical Software and Windows Vista - Ed.](#)

Can Somebody Give Me Some Help Here?

R: A Short Guide For The Uninitiated

By [Dr. Rich Herrington](#), Academic Computing and User Services, CITC

R is a collection of public domain software projects, most of which are licensed under the [GNU](#) license (see [CRAN](#)). Seasoned R users, in describing the R system to the uninitiated, usually DO NOT describe R as a "statistical package", but rather as a programming and development environment for statistical and graphical data analysis. This declaration can have a rather "off-putting" sound to it; however, to characterize the R project (and related projects, e.g. [Omega Project](#)) as merely a "stat" package is at best an understatement, and at worst a gross mischaracterization. R is NOT just an amazing collection of modern data analysis and statistical modeling tools - it is something more. The R project may have started as a "[stone soup](#)" collaborative, but it has in more than 10 years time grown in both size, and respectability within the communities of both academic and practicing statisticians and computational scientists. For example, a number of inter-disciplinary conferences have been organized to encourage the innovative use of R in training and research in the computational sciences and, in general, various scientific disciplines. Two notable conferences are the [UseR!](#) and [DSC conferences](#). Perhaps, to underscore the importance of this initiative from the perspective of a wider audience, it is interesting to point out that the [UseR! 2006](#) was in part sponsored by such notable corporate entities as: American Airlines, Merck, Wiley, Springer, Taylor and Francis (among others), and (in my mind) most importantly the [American Statistical Association](#). To quote the UseR! 2006 website: "...the program will focus on:

- R as the `lingua franca' of data analysis and statistical computing.
- providing a platform for R users to discuss and exchange ideas how R can be used to do statistical computations, data analysis, visualization and exciting applications in various fields.



Subscribe to
Benchmarks
Online

- giving an overview of the new features of the rapidly evolving R project....."

Surely, part of the excitement regarding R, is that the development of R has NOT only focused on the theoretical and methodological aspects of data analysis and graphical visualization. In other words, R is NOT just a "[statistical package](#)". R bears a strong resemblance to a scripting or programming language that allows for operating system level activities. For example, R could be thought of as having a strong resemblance in its capabilities to such indispensable operating system tools such as [Perl](#) or [Bash](#). The [R development team](#) (and thousands of contributors worldwide) have reused much of the GNU utilities and applications to great advantage. Here is a listing of a few notable examples:

- The R system provides [connections to GNU databases](#) such as [MySQL](#) and [PostgreSQL](#) as well as commercial databases, e.g. [Oracle](#).
- R allows for the capability of "[imbedding](#)" [Perl](#), [Python](#) and Java within the R language (and vice-versa) - that is, Perl, Python, and Java function calls from within an R script allow for an exchange of functionality and data between the foreign language environments and R (for Java see [SJava](#)).
- [APIs](#) exist for using [window widget constructors within R](#), based on Java, [Tcl-Tk](#), [Gtk](#) and others - useful for building menu driven user environments.
- Many utilities exist within R for manipulating [HTML and XML](#) based documents. The package "[R2HTML](#)" converts R objects to HTML output.
- An R client session can use the operating system [TCP/IP](#) stack and act as a HTTP client in both writing and receiving information (e.g. [GET and POST](#)) on an TCP/IP information network - R uses utilities such as "[wget](#)" and "[cURL](#)" (e.g. see R package "[Rcurl](#)"; for discussion and examples see the [JSS](#) paper [submitted by Duncan Lang](#)).
- [R can be utilized as a "CGI"](#) scripting language for populating web pages with dynamic content (also see [RApache](#) presentation; [RApache project page](#)).

This discussion of the versatility of R for system level activities might deemphasize the fact that the R system is primarily used for most researchers as a data analysis and statistical modeling tool. Although, in previous columns, I have emphasized the advantages in using [R in an educational setting](#) for teaching elementary and advanced statistics. There are far more [user contributed packages](#) available covering the statistical and graphical modeling of data (there are over a thousand user contributed packages available on the CRAN website). [The "CRAN Task View"](#) of the package content on CRAN gives a thematic view of the available packages and should help in locating information on packages that are of primary interest:

CRAN Task Views

| | |
|---------------------------------|---|
| Bayesian | Bayesian Inference |
| Cluster | Cluster Analysis & Finite Mixture Models |
| Econometrics | Computational Econometrics |
| Environmetrics | Analysis of ecological and environmental data |
| Finance | Empirical Finance |
| Genetics | Statistical Genetics |
| Graphics | Graphic Displays & Dynamic Graphics & Graphic Devices & Visualization |
| MachineLearning | Machine Learning & Statistical Learning |
| Multivariate | Multivariate Statistics |
| SocialSciences | Statistics for the Social Sciences |
| Spatial | Analysis of Spatial Data |
| gR | gRaphical models in R |

To automatically install these views, the `ctv` package needs to be installed, e.g., via `install.packages("ctv")` and then the views can be installed via `install.views` (after loading `ctv`), e.g., `install.views("Econometrics")`

Searching Web Forums For Help Using R

Q: What About R and Windows Vista?

R has a very well developed help system which includes both windows compiled help files ([CHM](#)) and [browser based HTML help files](#). Additionally, many [high quality manuals](#) are provided. The user community contributes through [active forums](#) and [user contributed documents](#) and tutorials. A number of websites feature Google searching of the archives of the many available forum websites (e.g. [Jonathon Baron's website](#)). With so many packages, so much functionality, and so many websites devoted to R, how is one to find help and direction that one needs during an ongoing session? Here, we will demonstrate the use of a function within R that capitalizes on R's ability to function as a web client. In an active R session we search within the R help for the function "[RSiteSearch](#)":

```

R Console
File Edit Misc Packages Help
> help(RSiteSearch)
Help for 'RSiteSearch' is shown in the browser
> RSiteSearch("Windows Vista")
A search query has been submitted to http://search.r-project.org
The results page should open in your browser shortly
> █

```

First we use the help function to look for the details on the ["RsiteSearch" function](#). Our question concerns any discussions that might be related to the use of [Windows Vista](#) with R. Searching on **"Windows Vista"** opens the default web browser and returns search results that are graded according to the degree of match with your query:

R Site Search

Note: more than two search terms may fail.

Query: [\[How to search\]](#)

Display: Description: Sort:

Target:

- Functions
- Documents
- R-help 2002-
- Rhelp 1997-2001
- R-devel

Results:

References:

- **docs:** [Windows: 16] [Vista: 2] [TOTAL: 2]
- **functions:** [Windows: 294] [Vista: 0] [TOTAL: 0]
- **Rhelp02a:** [Windows (Too many documents hit. Ignored)] [Vista: 95] [TOTAL: 95]

Total 97 documents matching your query.

1. [\[R\] Regarding Vista from christian.ritter at shell.com on 2007-03-28 \(stdin\)](#) (score: 59)
Author: christian.ritter at shell.com (christian.ritter) /> <meta name="Subject" content="2007-03-28" /> <style type="text/css
Date: Sun, 01 Apr 2007 08:54:23 -0500
[R] Regarding **Vista** This message: [Message body] [More options] Related messages:
Next message] [Alberto Monteiro: "[R] Help with
<http://finzi.psych.upenn.edu/R/Rhelp02a/archive/96296.html> (11,700 bytes)
2. [\[R\] Regarding Vista from Prof Brian Ripley on 2007-03-28 \(stdin\)](#) (score: 57)
Author: Prof Brian Ripley (ripley) /> <meta name="Subject" content="[R] Regara
/> <style type="text/css
Date: Sun, 01 Apr 2007 08:54:23 -0500
[R] Regarding **Vista** This message: [Message body] [More options] Related messages:
Schwartz: "[R] Import of a workspace into R" Pr
<http://finzi.psych.upenn.edu/R/Rhelp02a/archive/96293.html> (10,926 bytes)

Lastly, I would like point out that my experience with the R user community has been very positive. Folks are quick to give tips and advice....but remember, as with most computer

oriented user forums [RTFM](#) is mandatory if you wish to [avoid being flamed](#). Good luck in your forays into the world of R!.

Web References

- [R FAQ](#)
- [CRAN website](#)
- [Introduction to R](#)

Please note that information published in *Benchmarks Online* is likely to degrade over time, especially links to various Websites. To make sure you have the most current information on a specific topic, it may be best to search the UNT Website - <http://www.unt.edu> . You can also search *Benchmarks Online* - <http://www.unt.edu/benchmarks/archives/back.htm> as well as consult the UNT Helpdesk - <http://www.unt.edu/helpdesk/> Questions and comments should be directed to benchmarks@unt.edu

[Return to top](#)