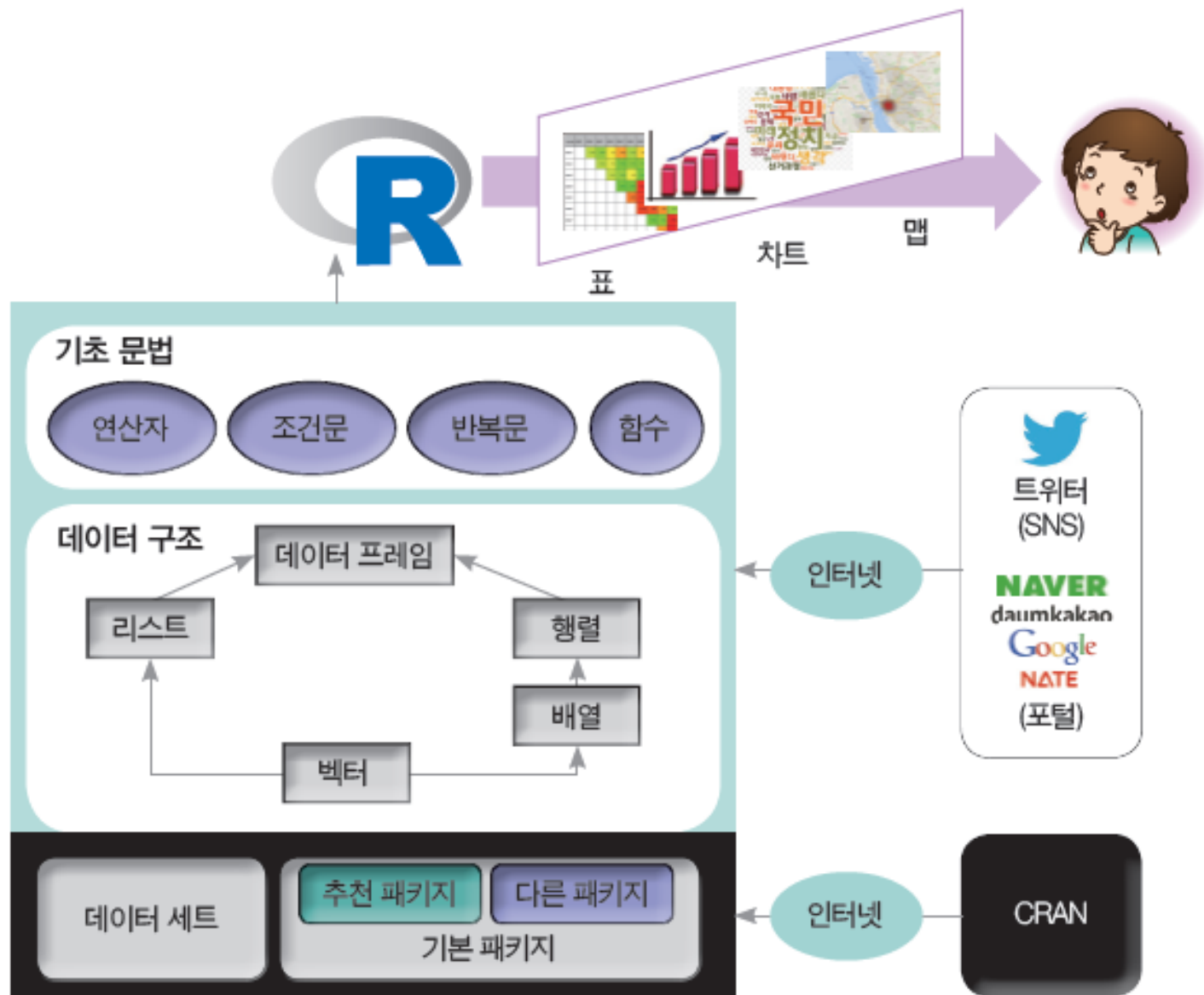


Getting Started

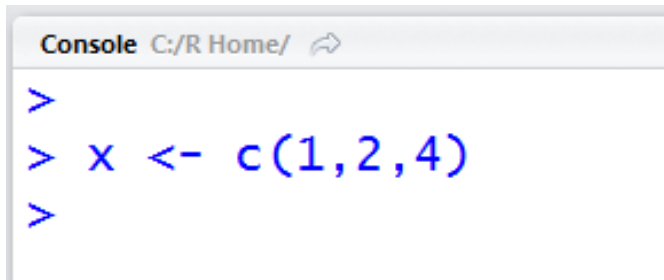
Bok, Jong Soon
javaexpert@nate.com
<https://github.com/swacademy/R>

R Usage Concepts



A First R Session

- Let's make a simple data set (in R parlance, a **vector**) consisting of the numbers **1**, **2**, and **4**, and name it **x**:

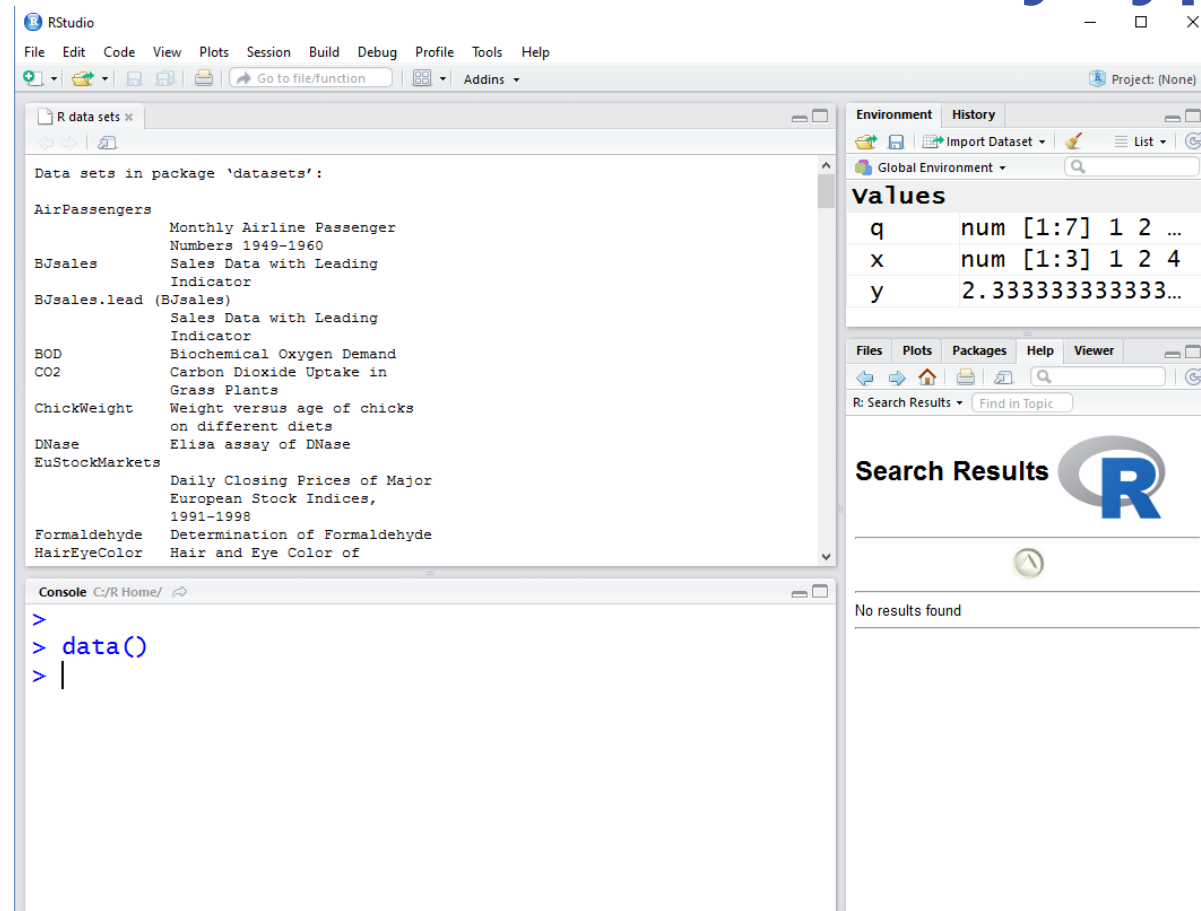


```
Console C:/R Home/ ↵  
>  
> x <- c(1,2,4)  
>
```

- The standard assignment operator in R is **<-**.
- You can also use **=**, but this is discouraged, as it does not work in some special situations.
- Note that there are no fixed types associated with variables.
- The **c** stands for concatenate.

A First R Session (Cont.)

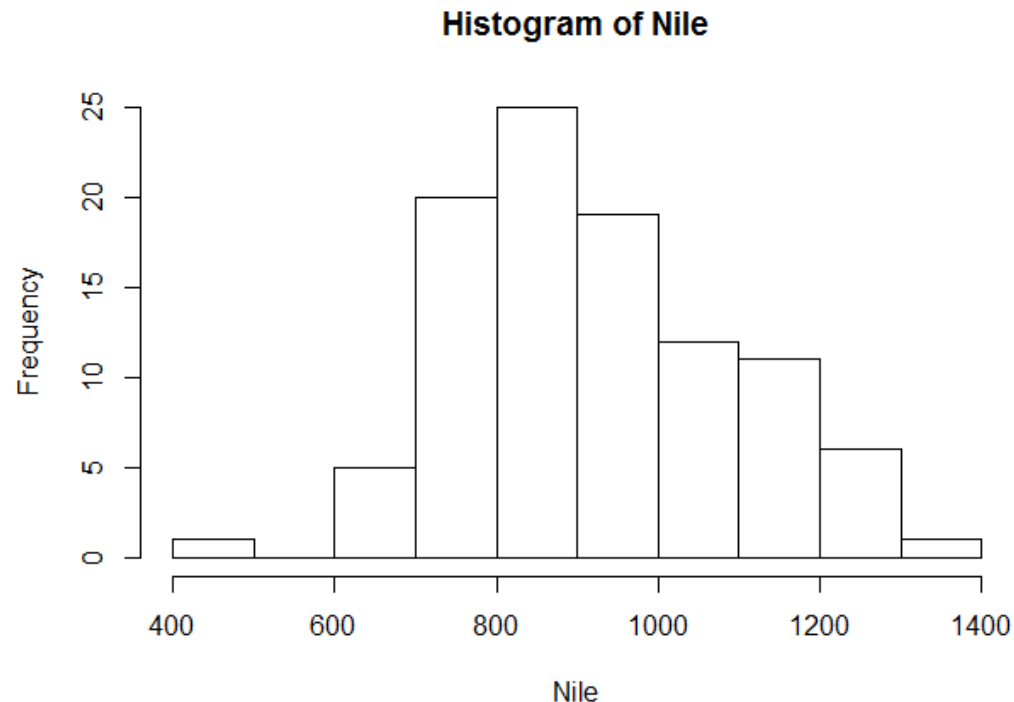
- Let's do something with one of R's internal data sets.
- You can get a list of these data sets by typing the following:



A First R Session (Cont.)

- One of the data sets is called **Nile** and contains data on the flow of the Nile River.
- Let's find the mean and standard deviation of this data set:

```
>  
> mean(Nile)  
[1] 919.35  
>  
> sd(Nile)  
[1] 169.2275  
>  
> hist(Nile)  
>
```



A First R Session (Cont.)

- Quit R by calling the `q()` function (or alternatively by pressing **CTRL-D** in Linux or **CMD-D** on a Mac):

```
Console C:/R Home/ ↵
>
>
> q()
Save workspace image to C:/R Home/.RData? [y/n]: |
```

In other hand...

Command Prompt

```
C:\WR Home>copy con test.R  
# My first program in R Programming  
myString <- "Hello, World!"
```

```
print( myString )
```

```
^Z
```

```
1 file(s) copied.
```

```
C:\WR Home>"C:\WProgram Files\WRWR-3.4.1\bin\WRScript.exe" test.R  
[1] "Hello, World!"
```

```
C:\WR Home>_
```

Introduction to Functions

- As in most programming languages, the heart of R programming consists of writing functions.
- A function is a group of instructions.

```
>
> # counts the number of odd intergers in x
> oddcount <- function(x){
+   k <- 0    # assign 0 to k
+   for(n in x) {
+     if(n %% 2 == 1) k <- k + 1  #%% is the modulo operator
+   }
+   return(k)
+ }
>
> oddcount(c(1,3,5))
[1] 3
>
> oddcount(c(1,2,3,7,9))
[1] 4
> |
```


Variable Scope

- A variable that is visible only within a function body is said to be *local* to that function.
- In `oddcount()`, `k` and `n` are *local* variables.

```
>  
> oddcount(c(1,2,3,7,9))  
[1] 4  
>  
> n  
Error: object 'n' not found  
>  
>  
> z <- c(2,6,7)  
> oddcount(z)  
[1] 1  
>
```

Variable Scope (Cont.)

- Variables created outside functions are *global* and are available within functions as well.

```
>  
> f <- function(x) return(x + y)  
> y <- 3  
> f(5)  
[1] 8  
>
```

- Here **y** is a *global* variable.

Some Important R Data Structures

- R has a variety of data structures.
- We will sketch some of the most frequently used structures.
- Scalars
- Character Strings
- Matrices
- Lists
- Data Frames
- Classes

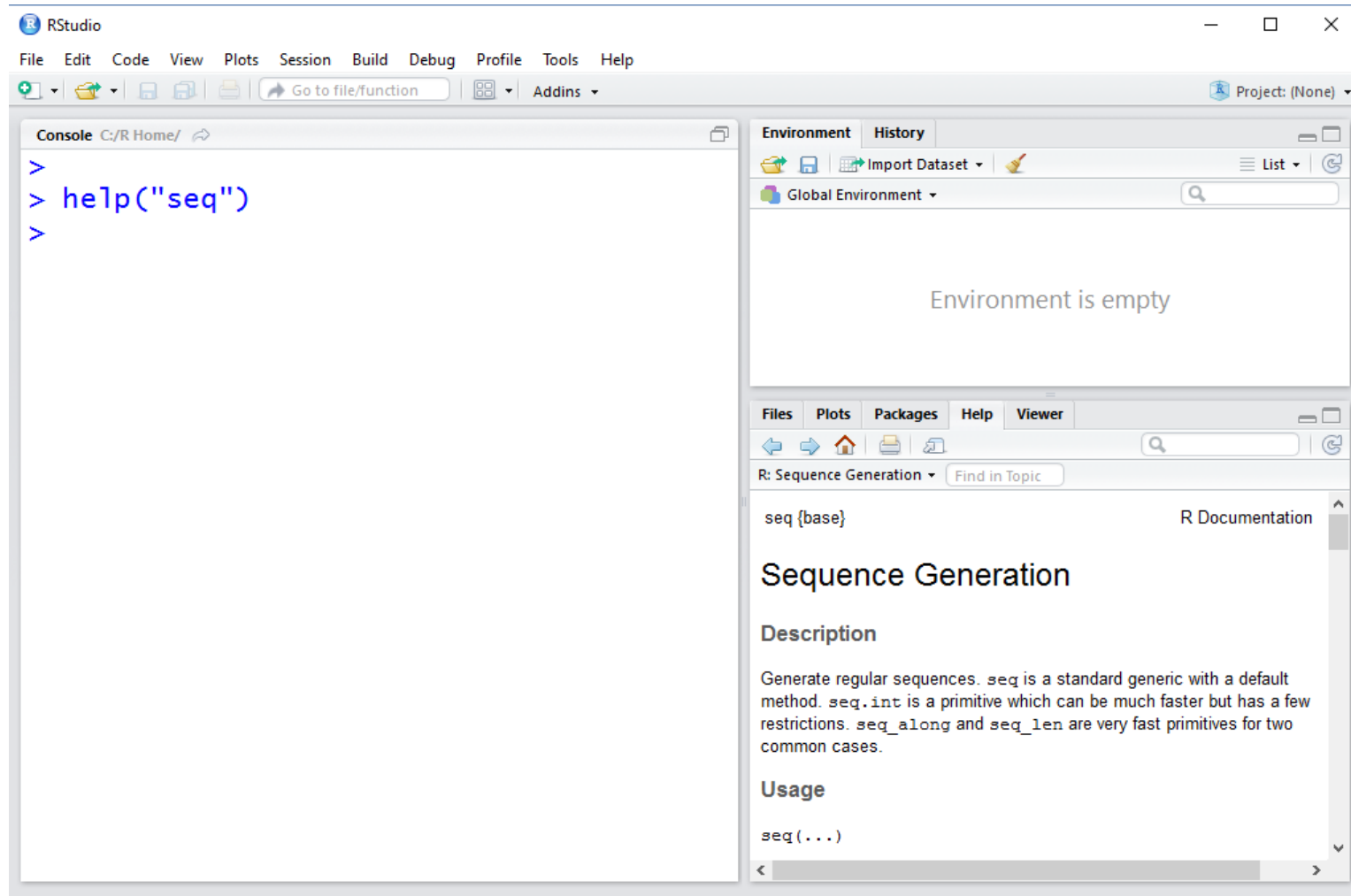
Comment in R

```
1
2  # 2017년 7월 6일
3
4  x <- 9    #x라는 변수에 9라는 정수형 스칼라값을 할당
5  y <- 5    #y라는 변수에 5라는 정수형 스칼라값을 할당
6
7  x         #x변수의 값 출력
8  y         #y변수의 값 출력
9
10
```

Getting Help

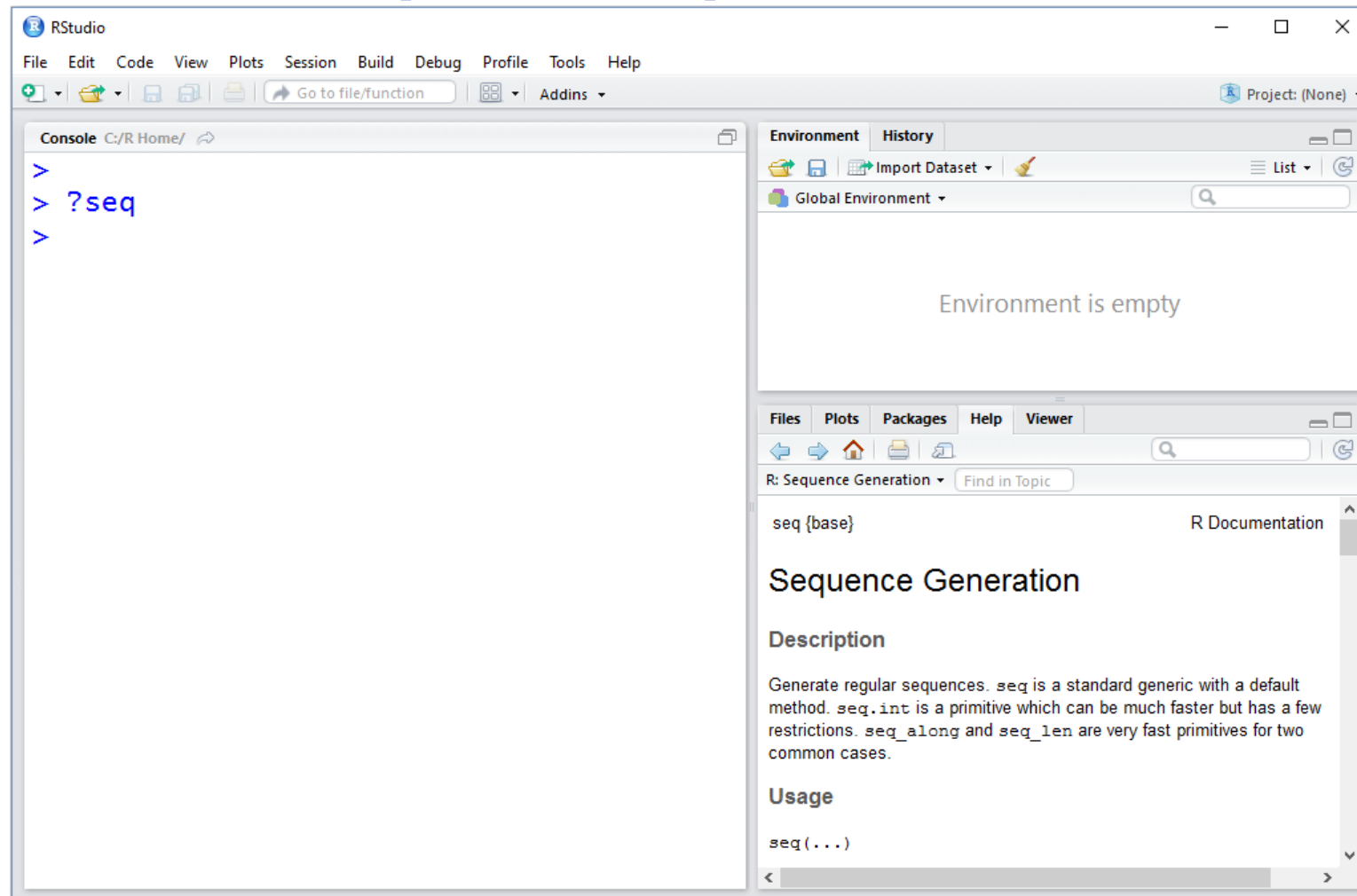
The help() Function

- To get online help, invoke **help()**.



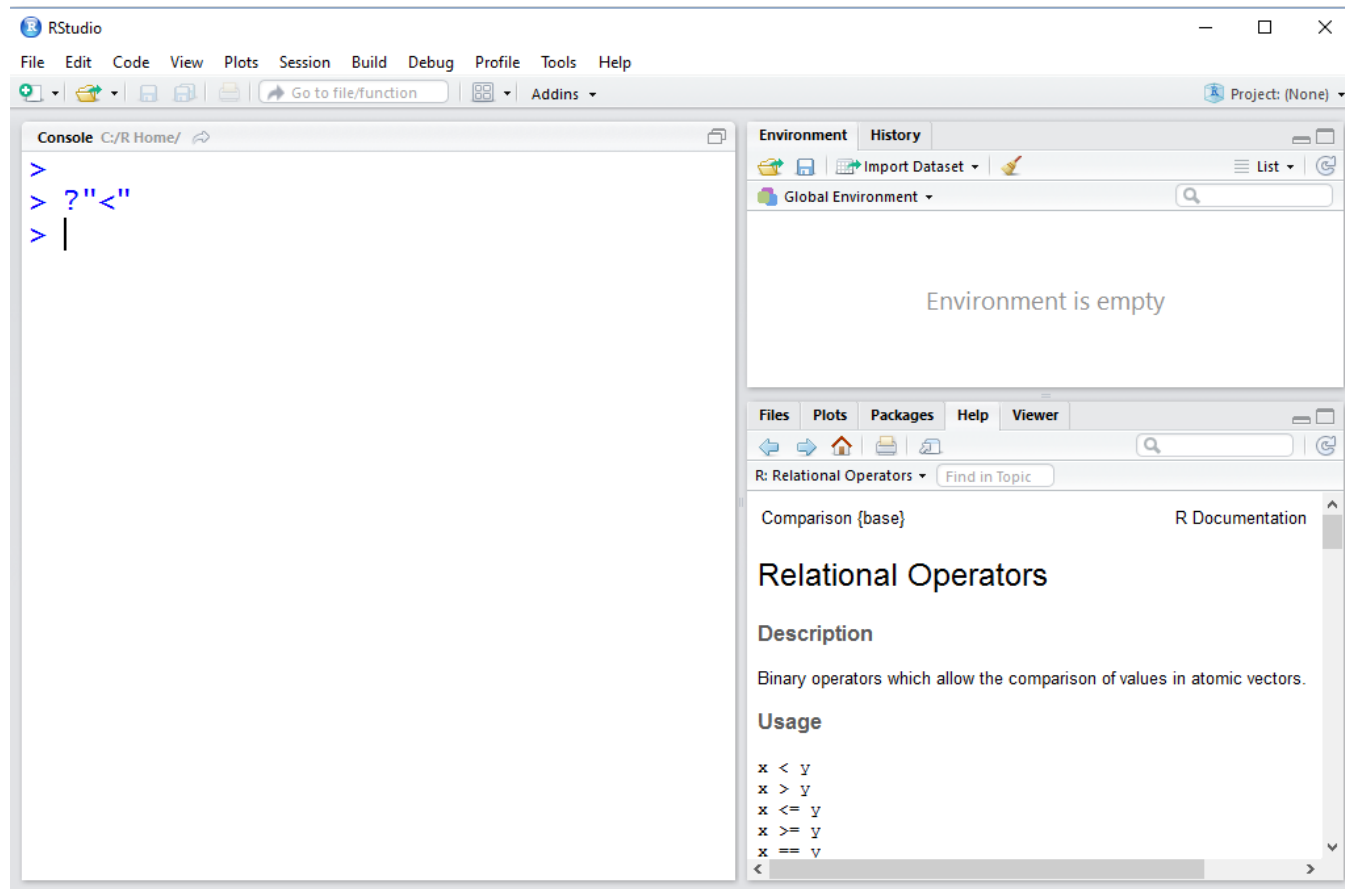
The help() Function

- The shortcut to help() is a question mark (?).



The help() Function

- Special characters and some reserved words must be quoted(" ") when used with the help() function.



The example() Function

- Each of the help entries comes with examples.
- One really nice feature of R is that the **example()** function will actually run those examples for you.

```
Console C:/R Home/
>
> example("seq")

seq> seq(0, 1, length.out = 11)
[1] 0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7
[9] 0.8 0.9 1.0

seq> seq(stats::rnorm(20)) # effectively 'along'
[1] 1 2 3 4 5 6 7 8 9 10 11
[12] 12 13 14 15 16 17 18 19 20

seq> seq(1, 9, by = 2) # matches 'end'
[1] 1 3 5 7 9

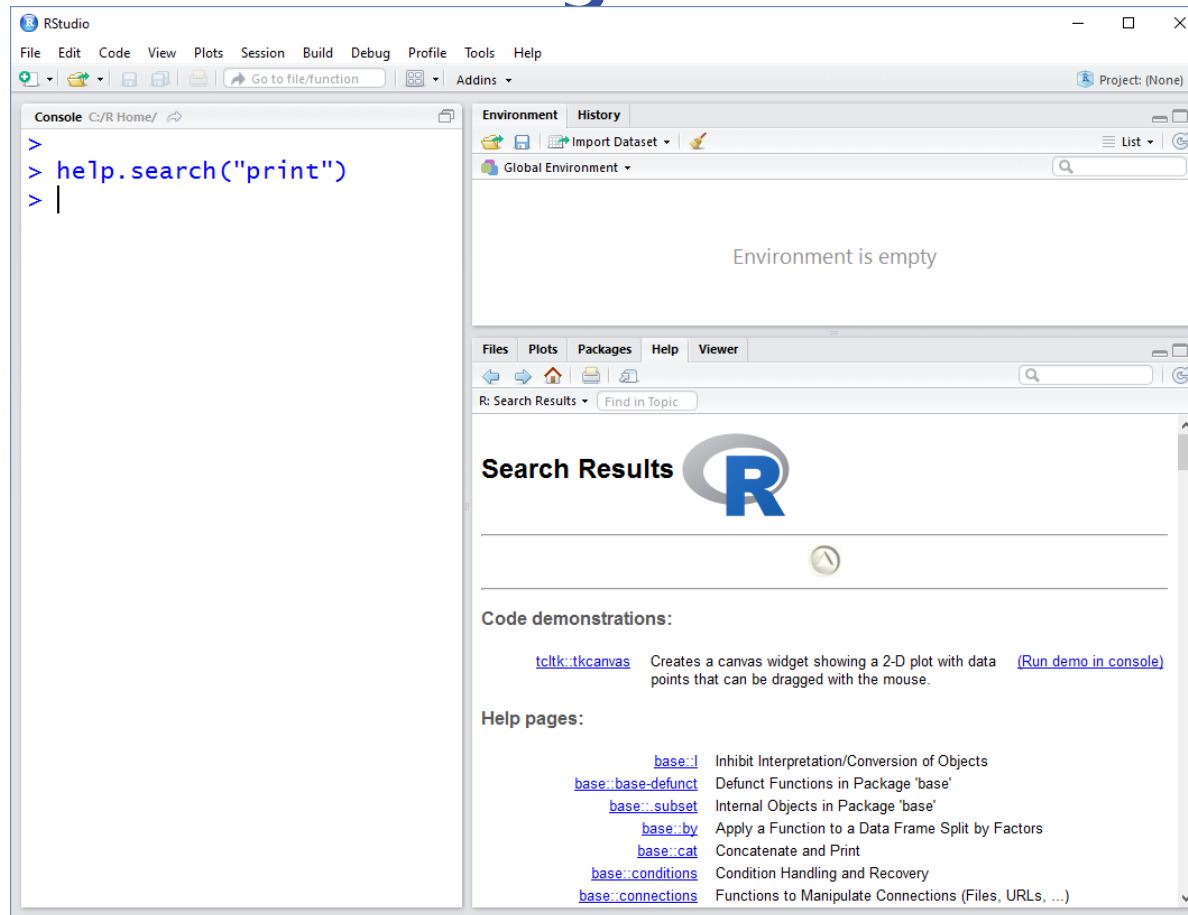
seq> seq(1, 9, by = pi) # stays below 'end'
[1] 1.000000 4.141593 7.283185

seq> seq(1, 6, by = 3)
[1] 1 4

seq> seq(1.575, 5.125, by = 0.05)
[1] 1.575 1.625 1.675 1.725 1.775
[6] 1.825 1.875 1.925 1.975 2.025
[11] 2.075 2.125 2.175 2.225 2.275
[16] 2.325 2.375 2.425 2.475 2.525
[21] 2.575 2.625 2.675 2.725 2.775
[26] 2.825 2.875 2.925 2.975 3.025
[31] 3.075 3.125 3.175 3.225 3.275
```

If You Don't Know Quite What You're Looking For

- You can use the function **help.search()** to do a Google-style search through R's documentation.



Help on the Internet

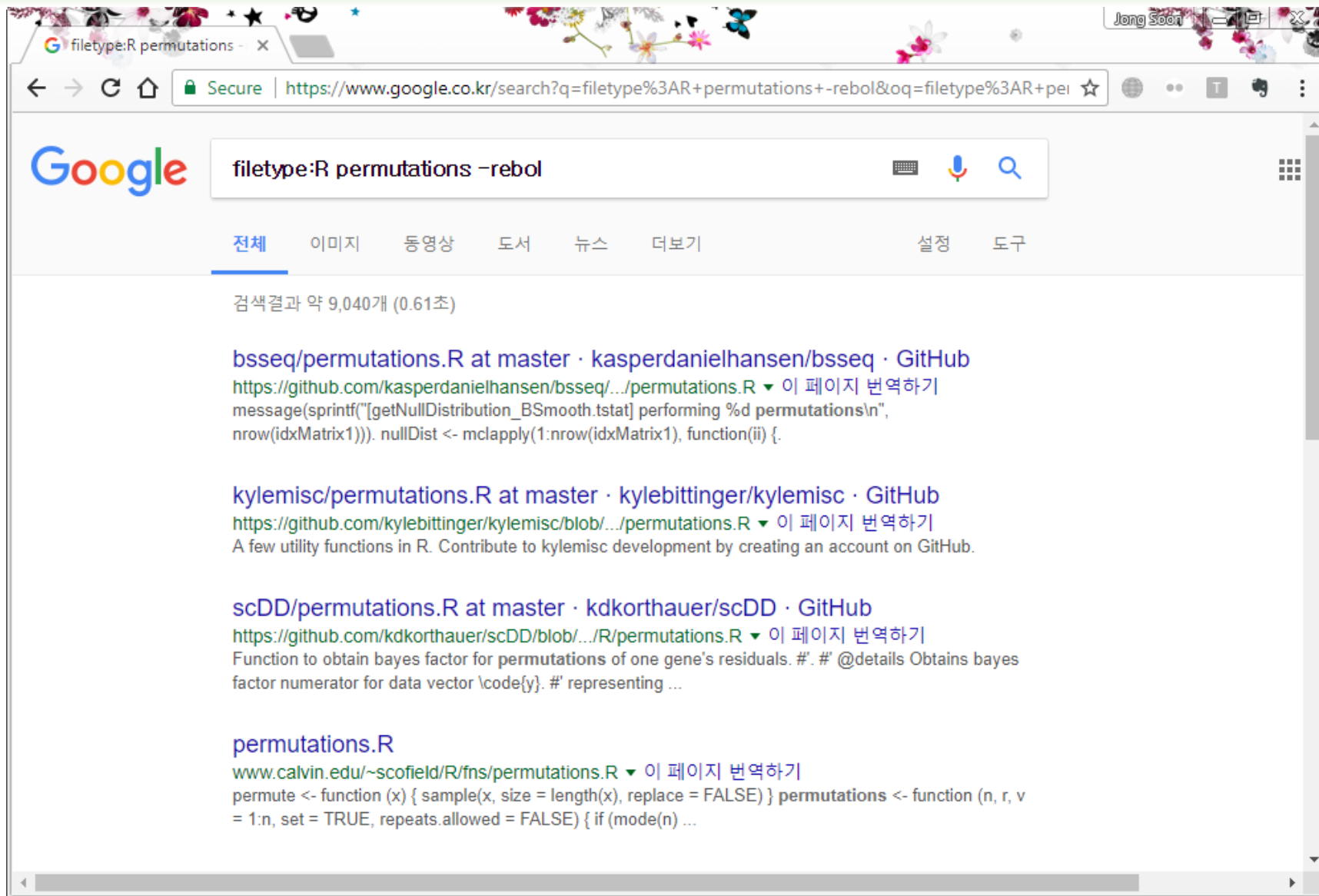
- The R Project's own manuals are available from the R home page, <http://www.r-project.org/>. → Click **Manuals**.
- Various R search engines are listed on the R home page. → Click **Search**.
- The **sos** package offers highly sophisticated searching of R materials.
 - See Appendix B for instructions on how to install R packages.
- RSeek search engine : <http://www.rseek.org/>.

Help on the Internet (Cont.)

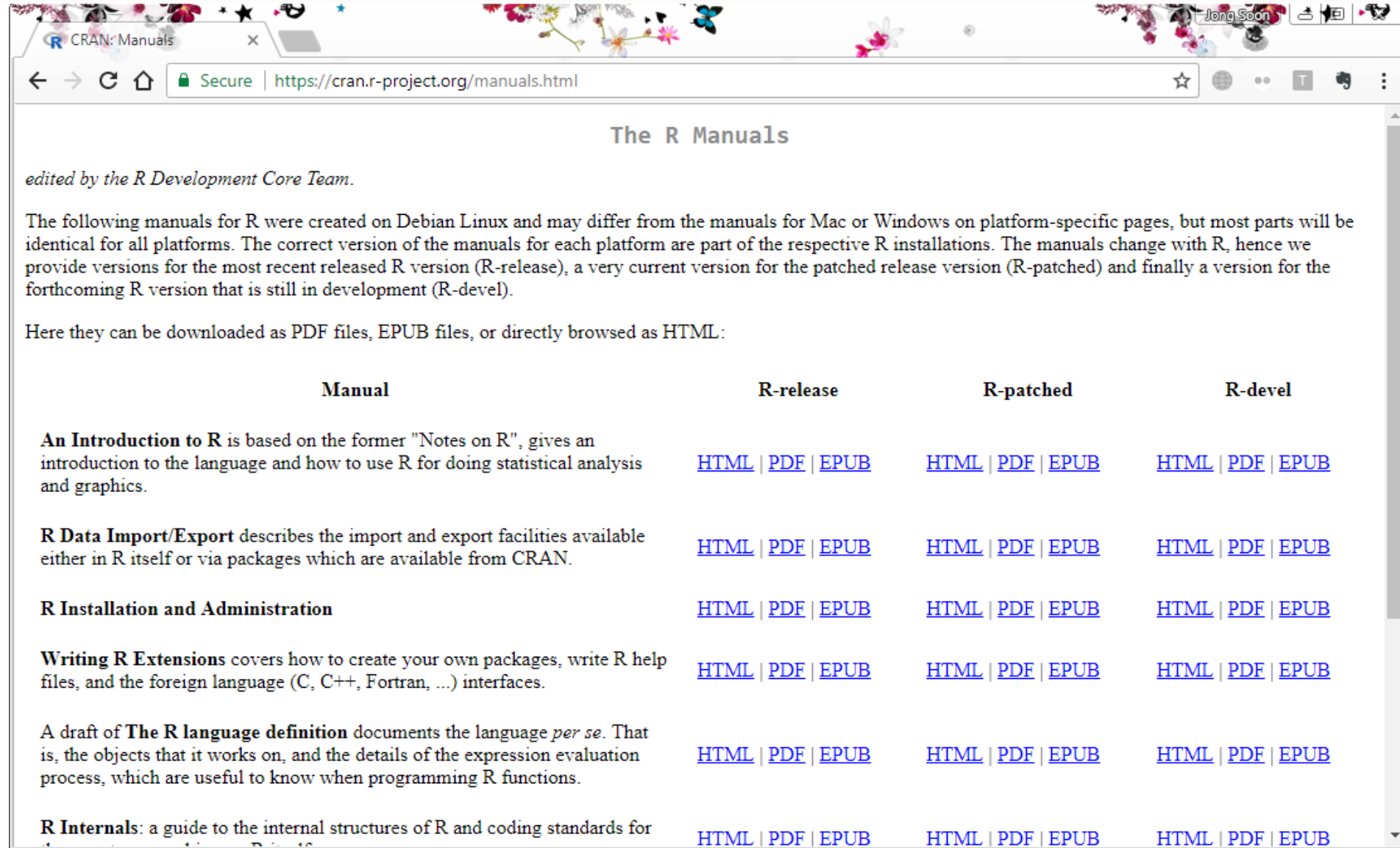
- Because of its single-letter name, R is difficult to search for using general purpose search engines such as Google.
- But there are tricks you can employ.
- One approach is to use Google's **filetype** criterion.

```
filetype:R permutations -rebol
```

Help on the Internet (Cont.)



Help on the Internet (Cont.)

A screenshot of a web browser displaying the CRAN Manuals page. The browser's address bar shows the URL 'https://cran.r-project.org/manuals.html'. The page title is 'The R Manuals', edited by the R Development Core Team. It provides an introduction to the manuals and lists several manual entries with links for different R versions (R-release, R-patched, R-devel) in HTML, PDF, and EPUB formats. The browser's taskbar at the top shows various icons and the name 'Jong Soon'.

CRAN: Manuals

Secure | https://cran.r-project.org/manuals.html

The R Manuals

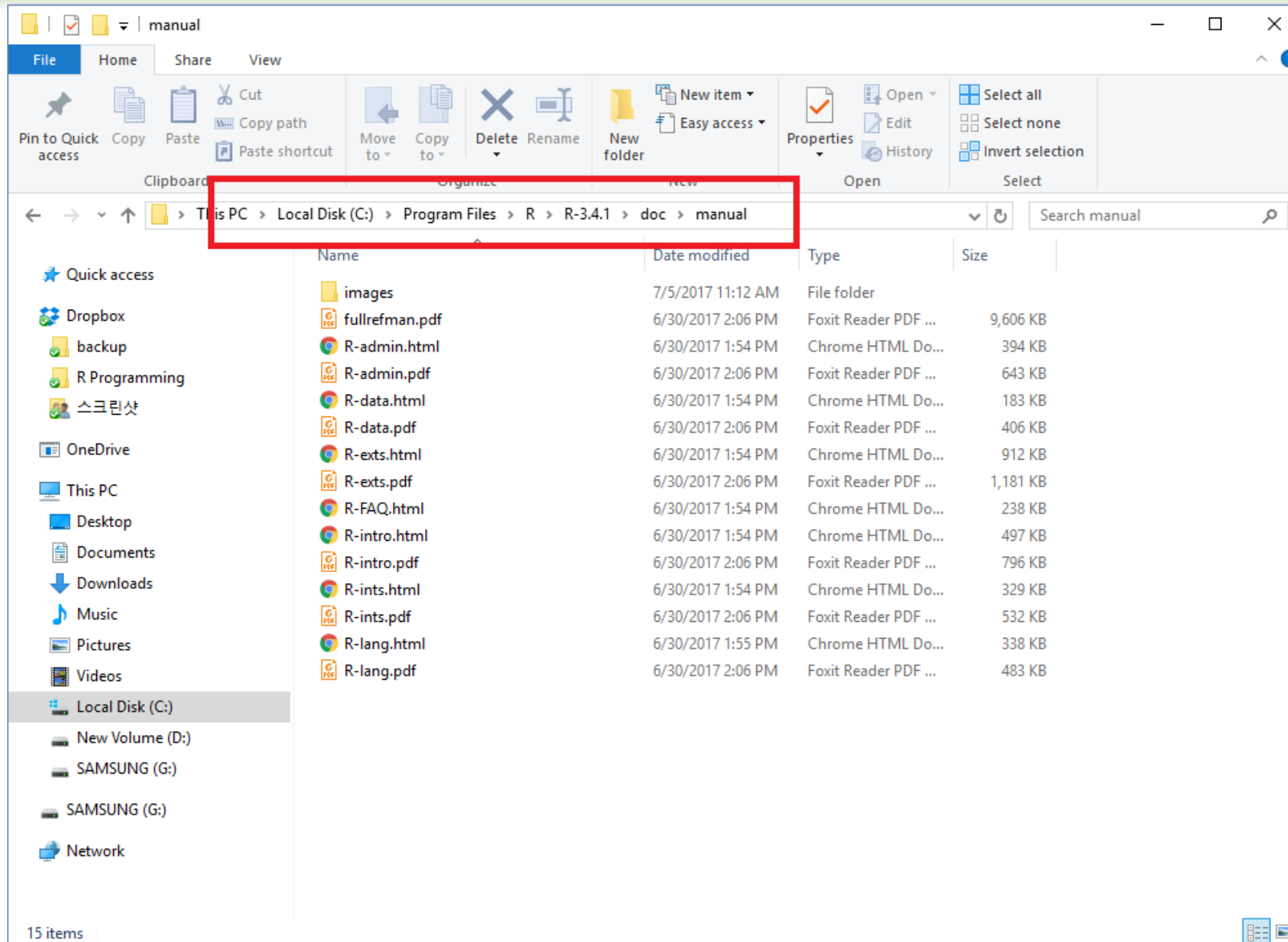
edited by the R Development Core Team.

The following manuals for R were created on Debian Linux and may differ from the manuals for Mac or Windows on platform-specific pages, but most parts will be identical for all platforms. The correct version of the manuals for each platform are part of the respective R installations. The manuals change with R, hence we provide versions for the most recent released R version (R-release), a very current version for the patched release version (R-patched) and finally a version for the forthcoming R version that is still in development (R-devel).

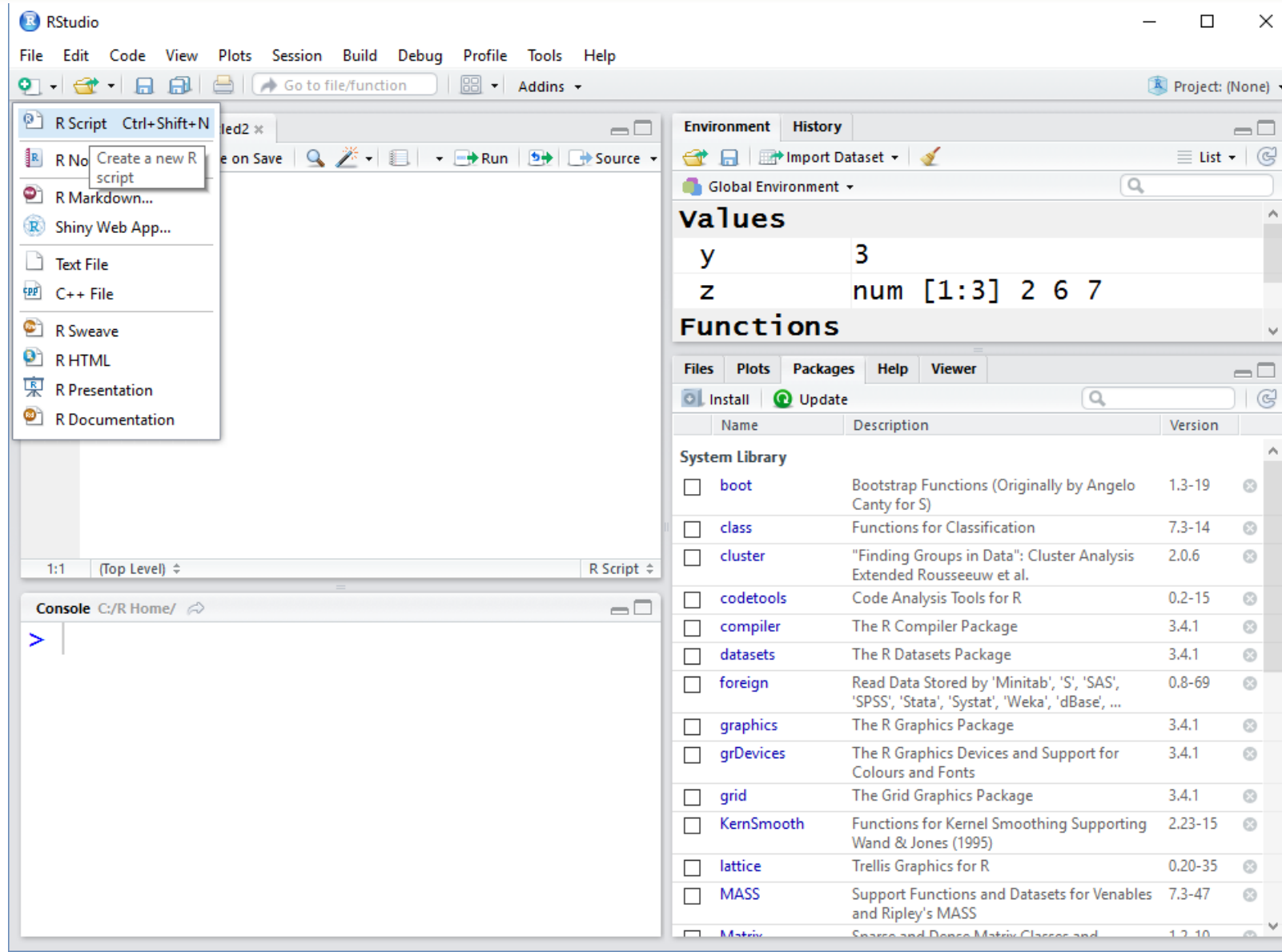
Here they can be downloaded as PDF files, EPUB files, or directly browsed as HTML:

Manual	R-release	R-patched	R-devel
An Introduction to R is based on the former "Notes on R", gives an introduction to the language and how to use R for doing statistical analysis and graphics.	HTML PDF EPUB	HTML PDF EPUB	HTML PDF EPUB
R Data Import/Export describes the import and export facilities available either in R itself or via packages which are available from CRAN.	HTML PDF EPUB	HTML PDF EPUB	HTML PDF EPUB
R Installation and Administration	HTML PDF EPUB	HTML PDF EPUB	HTML PDF EPUB
Writing R Extensions covers how to create your own packages, write R help files, and the foreign language (C, C++, Fortran, ...) interfaces.	HTML PDF EPUB	HTML PDF EPUB	HTML PDF EPUB
A draft of The R language definition documents the language <i>per se</i> . That is, the objects that it works on, and the details of the expression evaluation process, which are useful to know when programming R functions.	HTML PDF EPUB	HTML PDF EPUB	HTML PDF EPUB
R Internals : a guide to the internal structures of R and coding standards for	HTML PDF EPUB	HTML PDF EPUB	HTML PDF EPUB

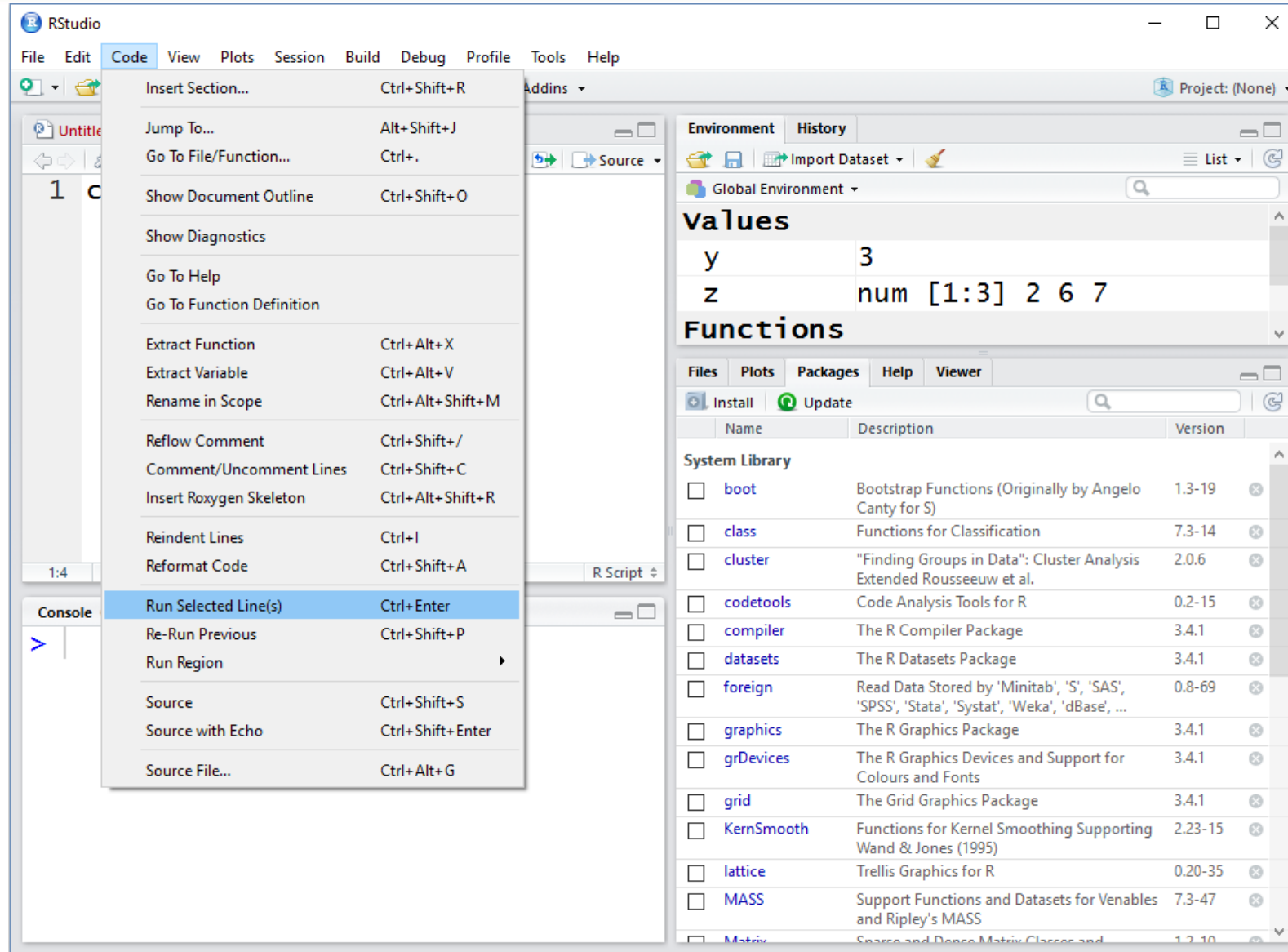
Local Document



R Script Window



R Script Window (Cont.)



R Script Window (Cont.)

The screenshot displays the RStudio application window. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu bar is a toolbar with icons for file operations and a search bar. The main workspace is divided into three panes:

- Source Editor (Left):** Contains three untitled files. The first file, 'Untitled1*', shows the following R code:

```
1 c(1,2,3)
2 |
```
- Environment (Right):** Displays the current environment. Under the 'Values' tab, it shows:
 - `y` with value `3`
 - `z` with value `num [1:3] 2 6 7`Under the 'Functions' tab, it shows a list of installed packages.
- Console (Bottom):** Shows the execution of the R code from the source editor:

```
> c(1,2,3)
[1] 1 2 3
> |
```

The bottom status bar indicates the current file is 'Untitled1*' and the session is at the 'Top Level'.

R Packages

- Are a collection of R functions, compiled code and sample data.
- Are stored under a directory called *library* in the R environment.
- By default, R installs a set of packages during installation.
- More packages are added later, when they are needed for some specific purpose.
- When start the R console, only the default packages are available by default.
- Other packages which are already installed have to be loaded explicitly to be used by the R program that is going to use them.
- All the packages available in R language are listed at R Packages.

R Packages (Cont.)

- **Check Available R Packages**

- Get library locations containing R packages

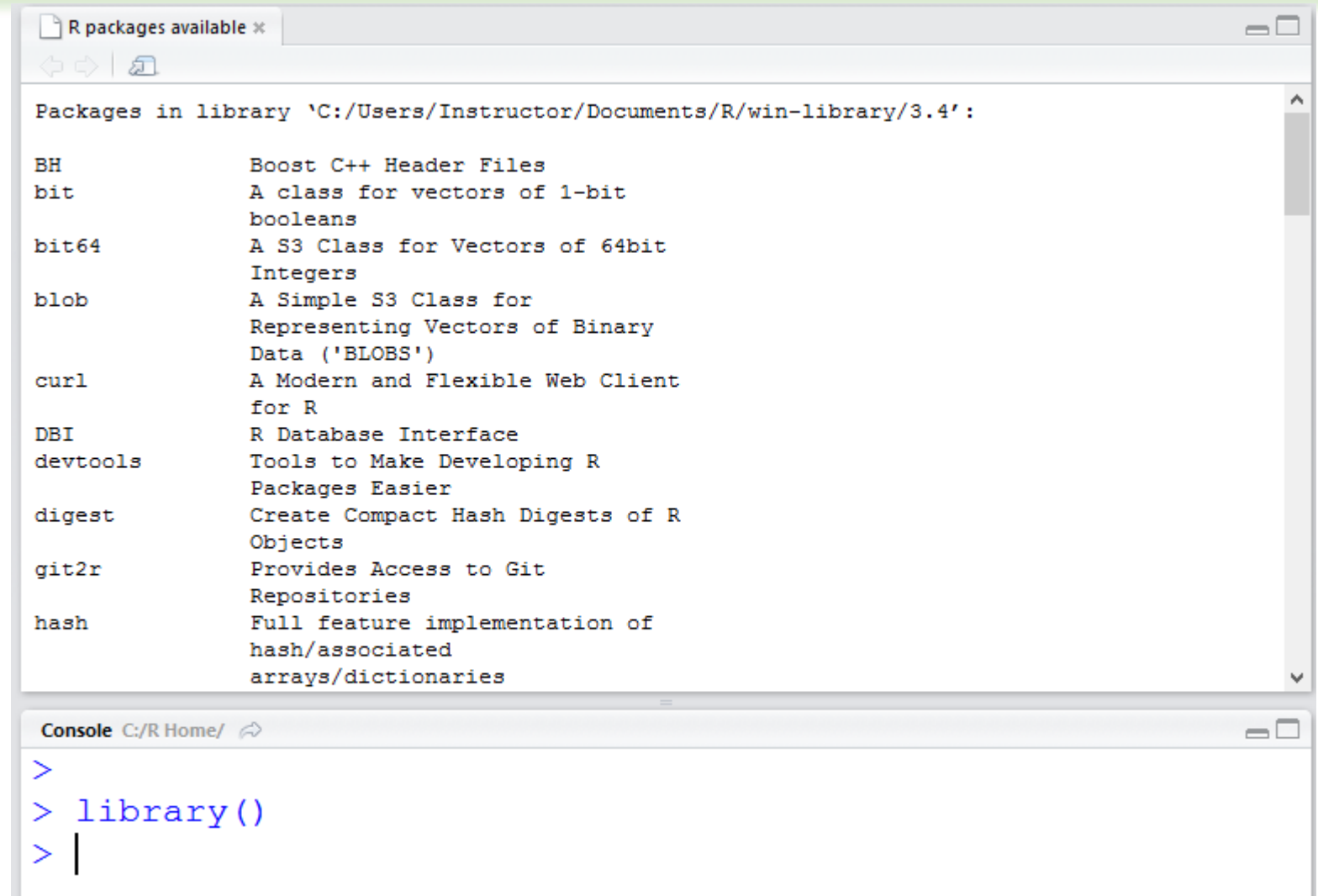
`.libPaths()`

```
Console C:/R Home/ ↵
> .libPaths()
[1] "C:/Users/Instructor/Documents/R/win-library/3.4"
[2] "C:/Program Files/R/R-3.4.1/library"
>
```

R Packages (Cont.)

- Get the list of all the packages installed

library()



The screenshot shows an R console window with two panes. The top pane, titled 'R packages available x', displays a list of installed packages and their descriptions. The bottom pane, titled 'Console C:/R Home/', shows the R prompt with the command `library()` entered.

```
R packages available x
Packages in library 'C:/Users/Instructor/Documents/R/win-library/3.4':

BH                Boost C++ Header Files
bit               A class for vectors of 1-bit
                 booleans
bit64             A S3 Class for Vectors of 64bit
                 Integers
blob              A Simple S3 Class for
                 Representing Vectors of Binary
                 Data ('BLOBS')
curl              A Modern and Flexible Web Client
                 for R
DBI               R Database Interface
devtools          Tools to Make Developing R
                 Packages Easier
digest            Create Compact Hash Digests of R
                 Objects
git2r             Provides Access to Git
                 Repositories
hash              Full feature implementation of
                 hash/associated
                 arrays/dictionaries

Console C:/R Home/
>
> library()
> |
```

R Packages (Cont.)

- Get all packages currently loaded in the R environment
search()

```
Console C:/R Home/ ↵
>
> search()
[1] ".GlobalEnv"          "tools:rstudio"
[3] "package:stats"        "package:graphics"
[5] "package:grDevices"    "package:utils"
[7] "package:datasets"     "package:methods"
[9] "Autoloads"            "package:base"
>
```

Package Installation & Usage

The screenshot displays the RStudio environment with the following components:

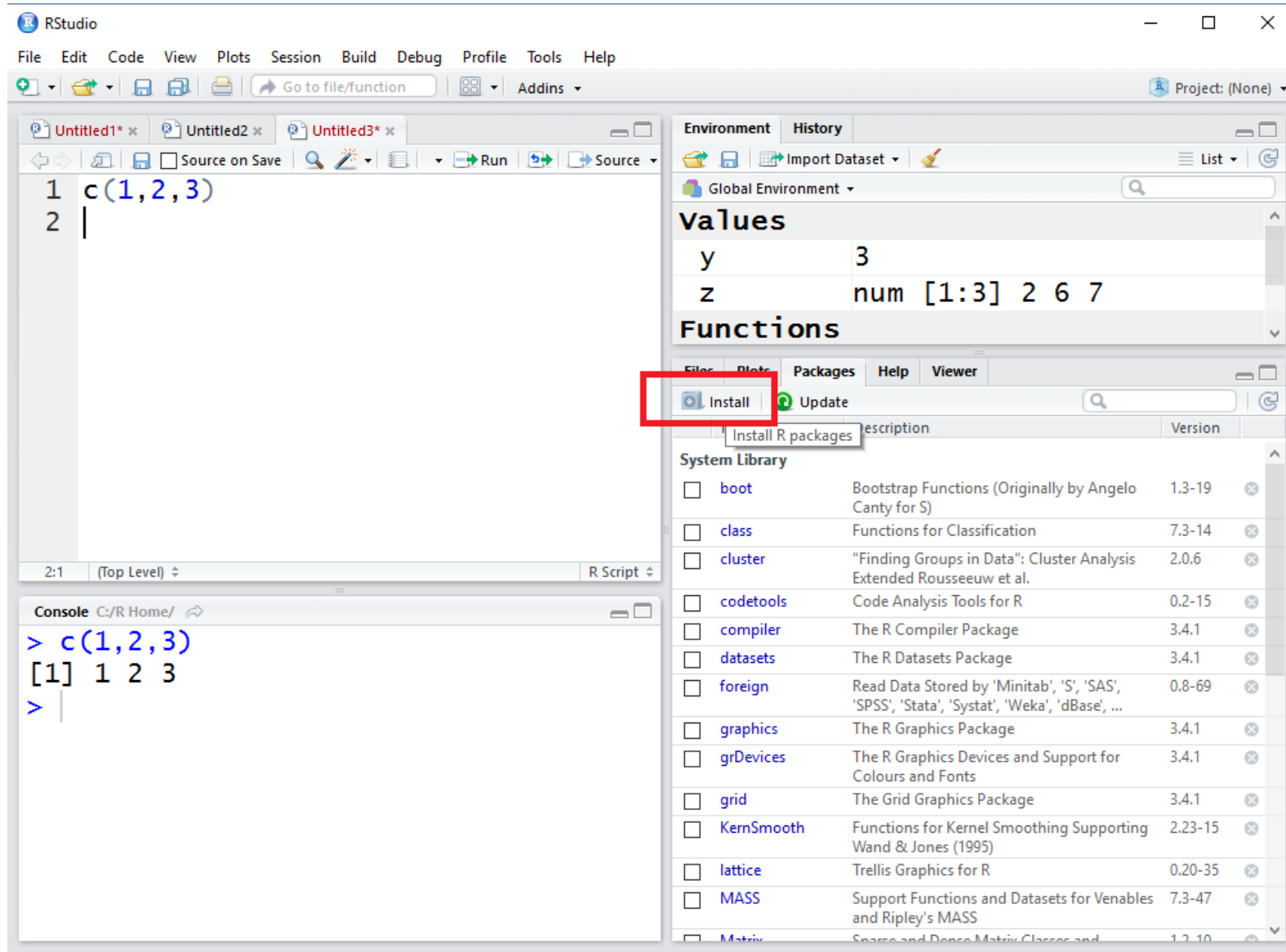
- Source Editor:** Contains three lines of R code:

```
1 c(1,2,3)
2
3 install.packages("KoNLP")
```
- Environment:** Shows variables `y` and `z`. Variable `y` has value `3`. Variable `z` has value `num [1:3] 2 6 7`.
- Console:** Shows the execution of the first line of code:

```
> c(1,2,3)
[1] 1 2 3
>
```
- Functions Panel:** Displays a list of installed and available packages. The 'System Library' section includes:

Name	Description	Version
<input type="checkbox"/> boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-19
<input type="checkbox"/> class	Functions for Classification	7.3-14
<input type="checkbox"/> cluster	"Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al.	2.0.6
<input type="checkbox"/> codetools	Code Analysis Tools for R	0.2-15
<input type="checkbox"/> compiler	The R Compiler Package	3.4.1
<input type="checkbox"/> datasets	The R Datasets Package	3.4.1
<input type="checkbox"/> foreign	Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', ...	0.8-69
<input type="checkbox"/> graphics	The R Graphics Package	3.4.1
<input type="checkbox"/> grDevices	The R Graphics Devices and Support for Colours and Fonts	3.4.1
<input type="checkbox"/> grid	The Grid Graphics Package	3.4.1
<input type="checkbox"/> KernSmooth	Functions for Kernel Smoothing Supporting Wand & Jones (1995)	2.23-15
<input type="checkbox"/> lattice	Trellis Graphics for R	0.20-35
<input type="checkbox"/> MASS	Support Functions and Datasets for Venables and Ripley's MASS	7.3-47
<input type="checkbox"/> Matrix	Sparse and Dense Matrix Classes and	1.2-10

Package Installation & Usage (Cont.)



The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains two lines of R code:

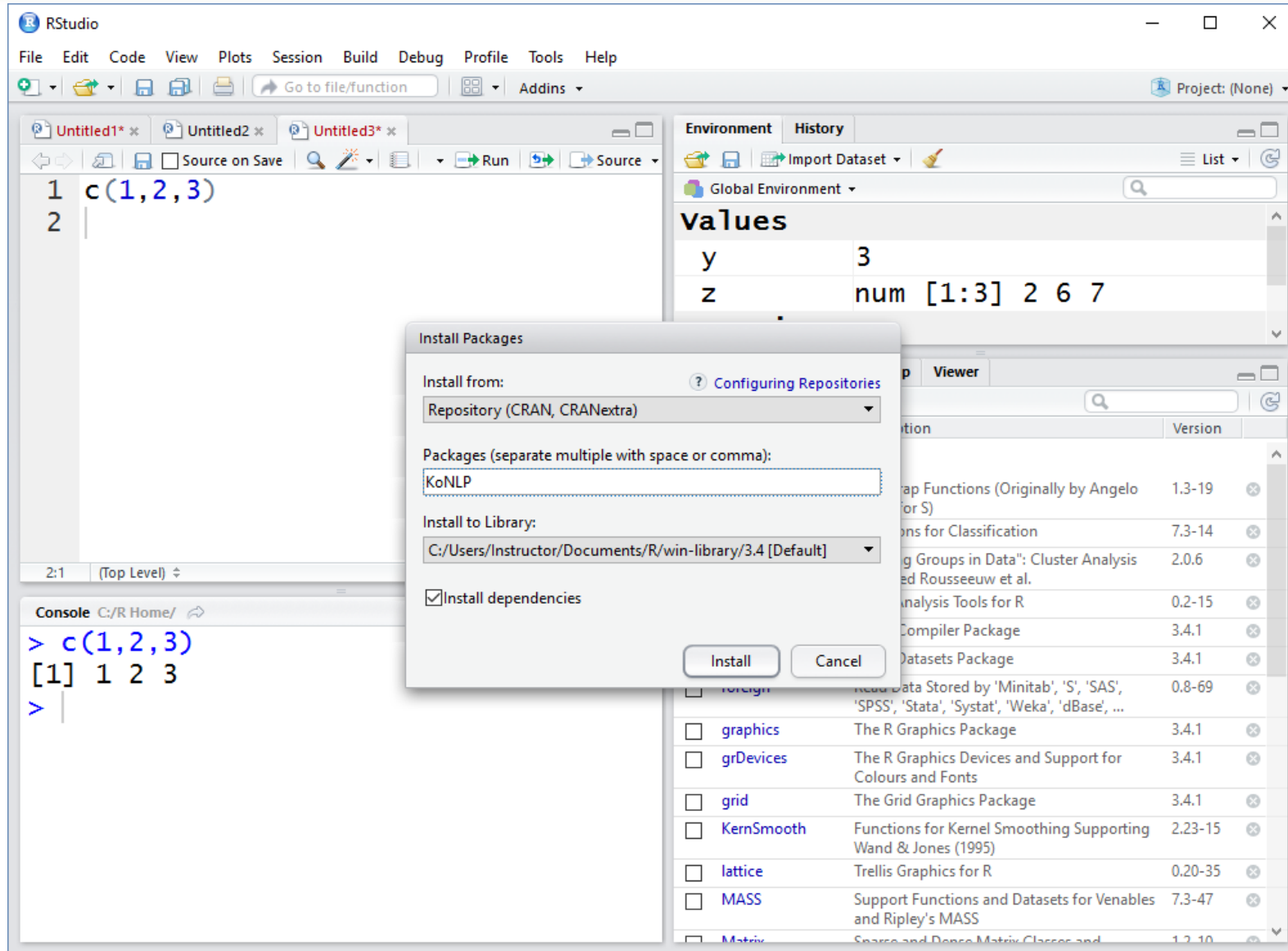
```
1 c(1,2,3)
2 |
```
- Console:** Shows the execution of the first line of code:

```
> c(1,2,3)
[1] 1 2 3
> |
```
- Environment Panel:** Displays the current environment with the following values:

Variable	Value
y	3
z	num [1:3] 2 6 7
- Packages Panel:** A red box highlights the **Install** button. Below it, a list of installed and available packages is shown:

Package	Description	Version
<input type="checkbox"/> boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-19
<input type="checkbox"/> class	Functions for Classification	7.3-14
<input type="checkbox"/> cluster	"Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al.	2.0.6
<input type="checkbox"/> codetools	Code Analysis Tools for R	0.2-15
<input type="checkbox"/> compiler	The R Compiler Package	3.4.1
<input type="checkbox"/> datasets	The R Datasets Package	3.4.1
<input type="checkbox"/> foreign	Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', ...	0.8-69
<input type="checkbox"/> graphics	The R Graphics Package	3.4.1
<input type="checkbox"/> grDevices	The R Graphics Devices and Support for Colours and Fonts	3.4.1
<input type="checkbox"/> grid	The Grid Graphics Package	3.4.1
<input type="checkbox"/> KernSmooth	Functions for Kernel Smoothing Supporting Wand & Jones (1995)	2.23-15
<input type="checkbox"/> lattice	Trellis Graphics for R	0.20-35
<input type="checkbox"/> MASS	Support Functions and Datasets for Venables and Ripley's MASS	7.3-47
<input type="checkbox"/> Matrix	Sparse and Dense Matrix Classes and	1.2-10

Package Installation & Usage (Cont.)



Package Installation & Usage (Cont.)

The screenshot displays the RStudio interface during the installation of a package. The main editor window shows the R code `c(1,2,3)` in the first line of `Untitled1.R`. The Environment pane on the right shows the global environment with variables `y` (value 3) and `z` (value `num [1:3] 2 6 7`). The Packages pane on the right lists installed and available packages, including `compiler`, `datasets`, `foreign`, `graphics`, `grDevices`, `grid`, `KernSmooth`, `lattice`, `MASS`, and `Matrix`. A progress dialog box is overlaid on the console, showing that 72% of the package has been downloaded from the URL `...s://cran.rstudio.com/bin/windows/contrib/3.4/BH_1.62.0-1.zip`. The console output at the bottom shows the progress of the download and the attempt to install the package from the CRAN mirror.

```
1 c(1,2,3)
2 |
```

72% downloaded

URL: ...s://cran.rstudio.com/bin/windows/contrib/3.4/BH_1.62.0-1.zip

```
Content type 'application/zip' length
h 3380737 bytes (3.2 MB)
downloaded 3.2 MB

trying URL 'https://cran.rstudio.com
/bin/windows/contrib/3.4/BH_1.62.0-1
.zip'
Content type 'application/zip' lengt
h 16150325 bytes (15.4 MB)
```

Name	Description	Version
compiler	The R Compiler Package	3.4.1
datasets	The R Datasets Package	3.4.1
foreign	Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', ...	0.8-69
graphics	The R Graphics Package	3.4.1
grDevices	The R Graphics Devices and Support for Colours and Fonts	3.4.1
grid	The Grid Graphics Package	3.4.1
KernSmooth	Functions for Kernel Smoothing Supporting Wand & Jones (1995)	2.23-15
lattice	Trellis Graphics for R	0.20-35
MASS	Support Functions and Datasets for Venables and Ripley's MASS	7.3-47
Matrix	Sparse and Dense Matrix Classes and	1.2-10

Package Installation & Usage (Cont.)

The screenshot shows the RStudio interface with the following components:

- Source Editor:** Contains two lines of R code:

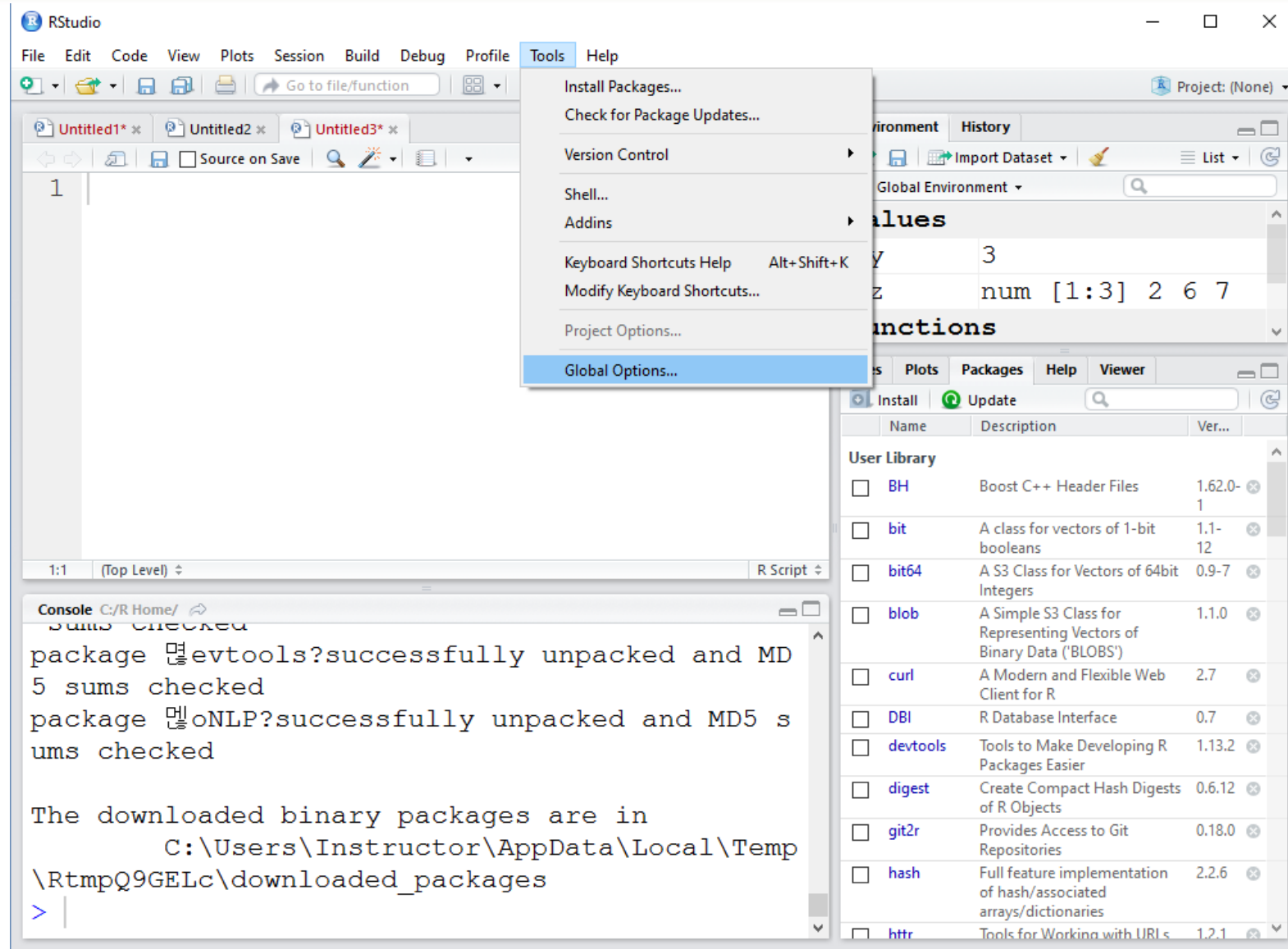
```
1 c(1,2,3)
2 |
```
- Console:** Displays the output of the installation process:

```
package 'evtools'?successfully unpacked and MD5 sums checked
package 'KoNLP'?successfully unpacked and MD5 sums checked

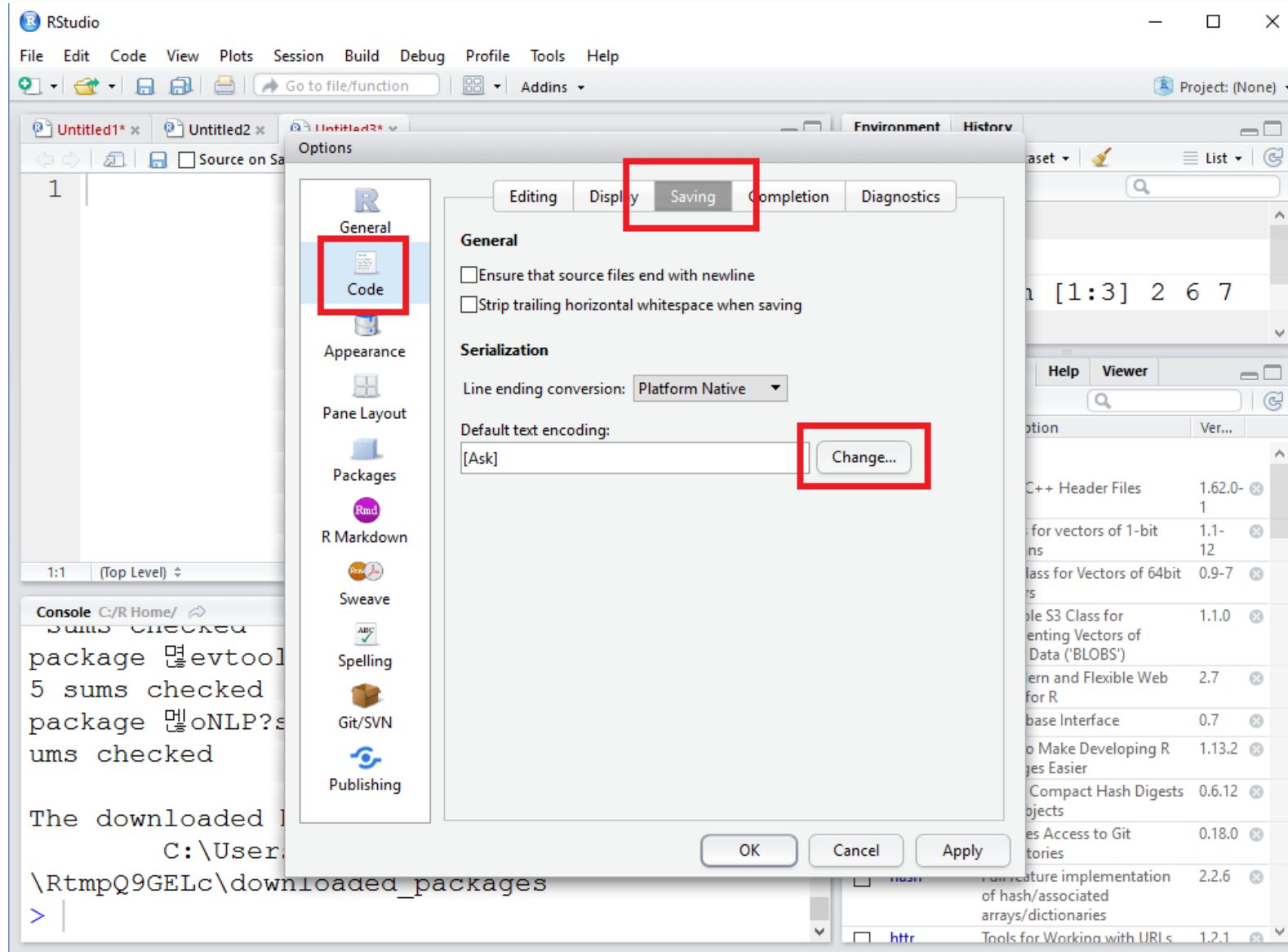
The downloaded binary packages are in
  c:\Users\Instructor\AppData\Local\Temp\RtmpQ9GELc\downloaded_packages
```
- Environment Panel:** Shows the current environment with variables `y` (value 3) and `z` (value `num [1:3] 2 6 7`).
- Packages Panel:** Lists installed and available packages. The `KoNLP` package is highlighted with a red box. The table below represents the data in this panel:

Name	Description	Ver...
<input type="checkbox"/> httr	Tools for Working with URLs and HTTP	1.2.1
<input type="checkbox"/> jsonlite	A Robust, High Performance JSON Parser and Generator for R	1.5
<input type="checkbox"/> KoNLP	Korean NLP Package	0.80.1
<input type="checkbox"/> magrittr	A Forward-Pipe Operator for R	1.5
<input type="checkbox"/> memoise	Memoisation of Functions	1.1.0
<input type="checkbox"/> mime	Map Filenames to MIME Types	0.5
<input type="checkbox"/> openssl	Toolkit for Encryption, Signatures and Certificates Based on OpenSSL	0.9.6
<input type="checkbox"/> pkgconfig	Private Configuration for 'R' Packages	2.0.1
<input type="checkbox"/> plogr	The 'plog' C++ Logging Library	0.1-1
<input type="checkbox"/> R6	Classes with Reference Semantics	2.2.2
<input type="checkbox"/> Rcpp	Seamless R and C++ Integration	0.12.11

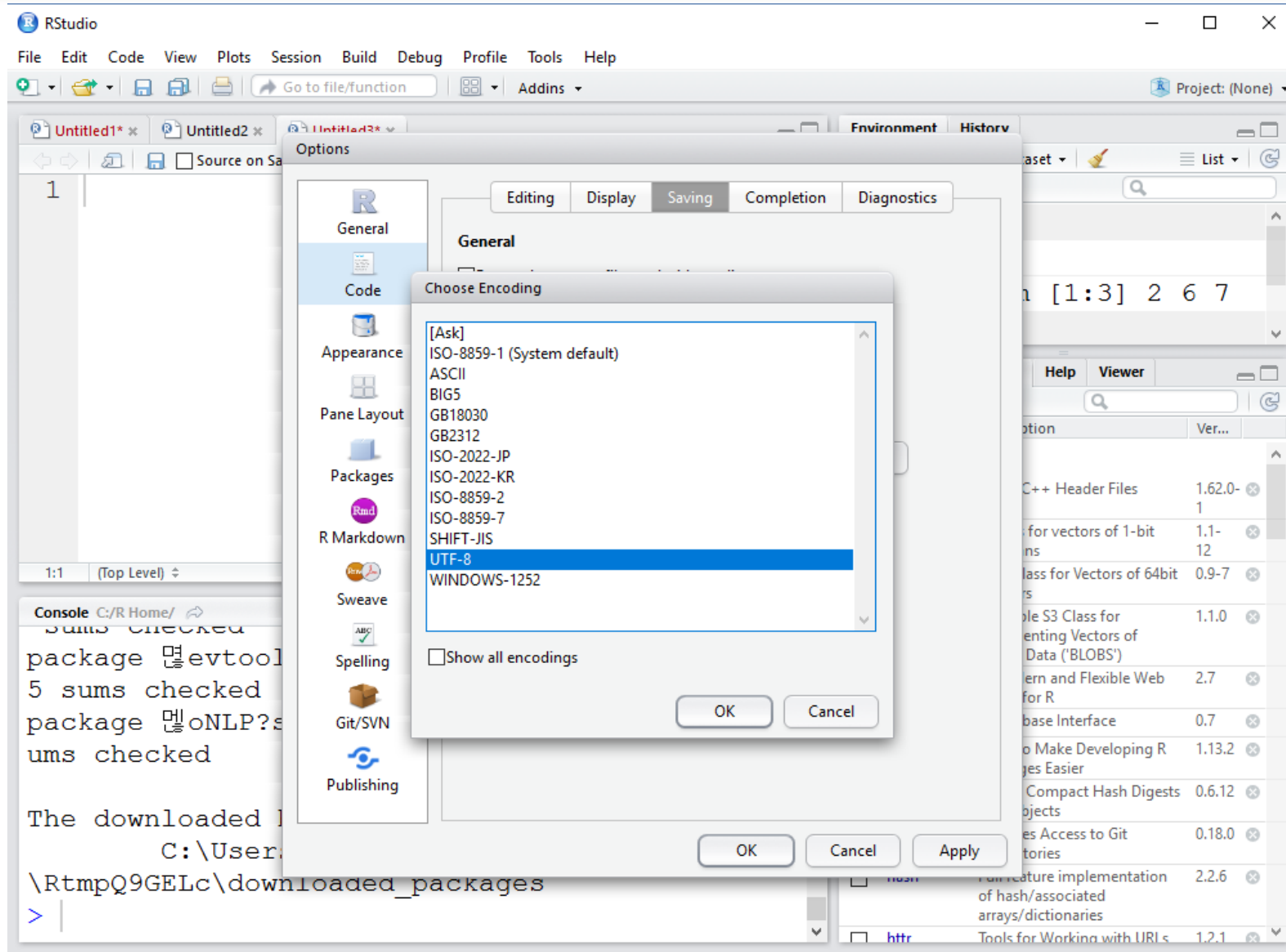
Package Installation & Usage (Cont.)



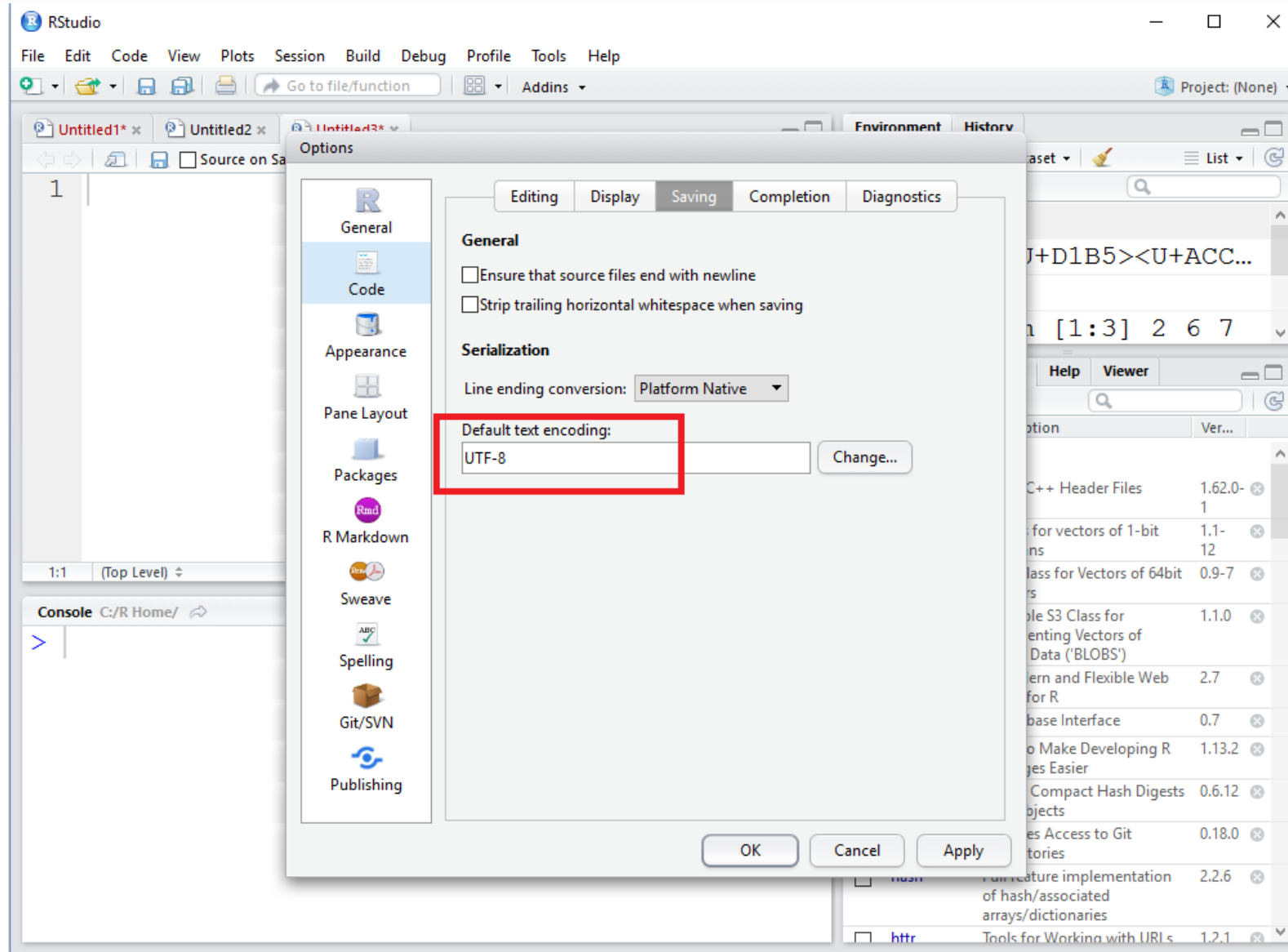
Package Installation & Usage (Cont.)



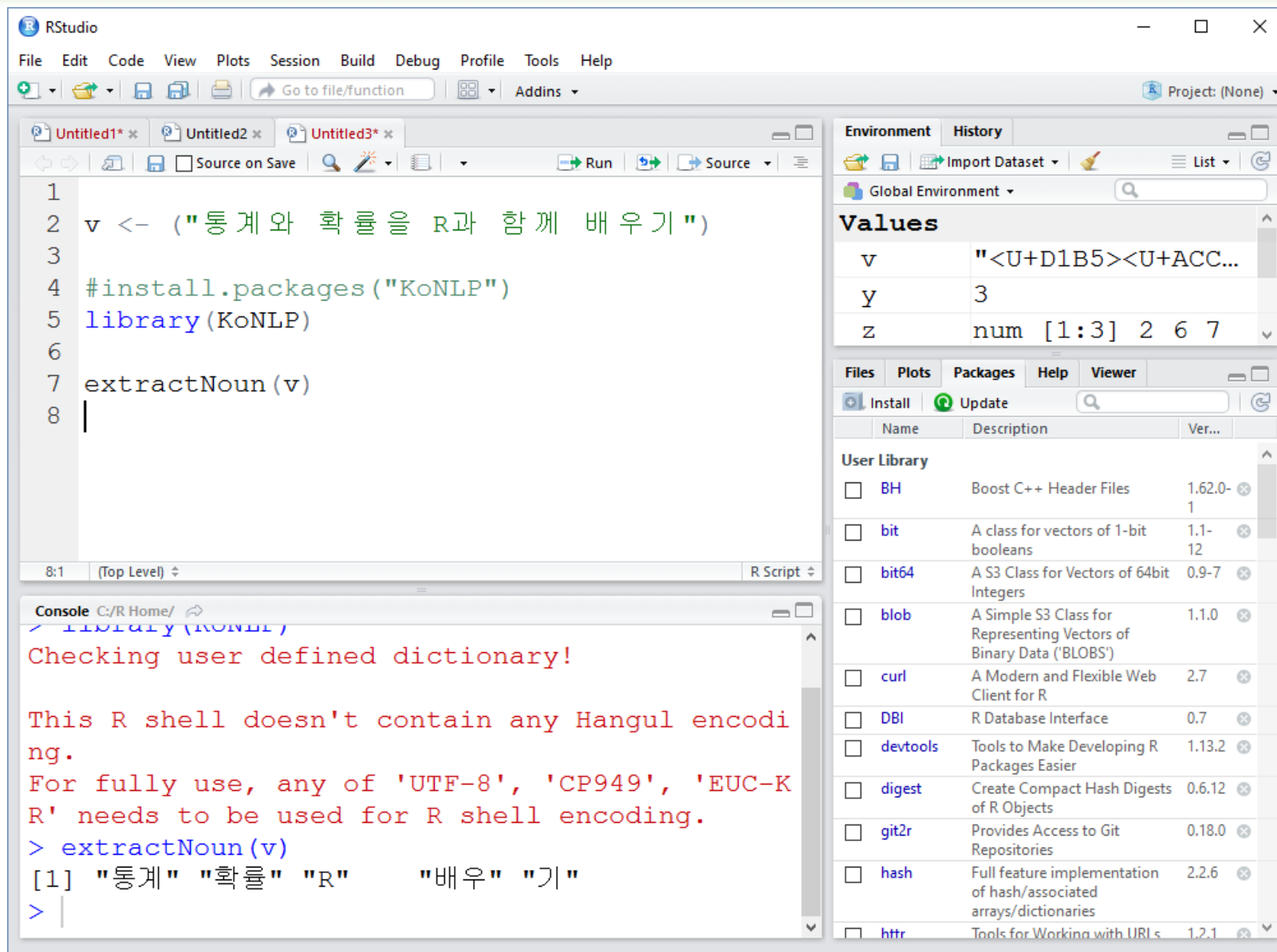
Package Installation & Usage (Cont.)



Package Installation & Usage (Cont.)



Package Installation & Usage (Cont.)



The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for installing and using the `KoNLP` package. The code defines a variable `v` with Korean text, installs `KoNLP`, loads it, and uses `extractNoun` to extract nouns from the text.
- Environment:** Shows the current environment with variables `v` (a character vector), `y` (a numeric value), and `z` (a numeric vector).
- Console:** Shows the output of the code execution, including a warning about the R shell's encoding and the results of `extractNoun`.
- Packages:** A list of installed and available packages, including `BH`, `bit`, `bit64`, `blob`, `curl`, `DBI`, `devtools`, `digest`, `git2r`, `hash`, and `httr`.

```
1 v <- ("통계와 확률을 R과 함께 배우기")
2
3
4 #install.packages("KoNLP")
5 library(KoNLP)
6
7 extractNoun(v)
8
```

Console Output:

```
> library(KoNLP)
Checking user defined dictionary!

This R shell doesn't contain any Hangul encoding.
For fully use, any of 'UTF-8', 'CP949', 'EUC-KR' needs to be used for R shell encoding.
> extractNoun(v)
[1] "통계" "확률" "R"      "배우" "기"
```

Name	Description	Ver...
<input type="checkbox"/> BH	Boost C++ Header Files	1.62.0-1
<input type="checkbox"/> bit	A class for vectors of 1-bit booleans	1.1-12
<input type="checkbox"/> bit64	A S3 Class for Vectors of 64bit Integers	0.9-7
<input type="checkbox"/> blob	A Simple S3 Class for Representing Vectors of Binary Data ('BLOBS')	1.1.0
<input type="checkbox"/> curl	A Modern and Flexible Web Client for R	2.7
<input type="checkbox"/> DBI	R Database Interface	0.7
<input type="checkbox"/> devtools	Tools to Make Developing R Packages Easier	1.13.2
<input type="checkbox"/> digest	Create Compact Hash Digests of R Objects	0.6.12
<input type="checkbox"/> git2r	Provides Access to Git Repositories	0.18.0
<input type="checkbox"/> hash	Full feature implementation of hash/associated arrays/dictionaries	2.2.6
<input type="checkbox"/> httr	Tools for Working with URLs	1.2.1

Package Remove

- To remove installed packages, do like below:

`remove.packages("package_name")`

