Using R for Data Analysis and Data Visualization

Instructor: Li, Han

- a calculator
- generate random variables/distributions, for example, normal distribution
- data visualization, that's, plot the data
- data analysis, for example, linear regression

Almost all the things related to data analysis!

Why it is named 

R is named partly after the first names of the two authors who created it: Ross Ihaka and Robert Gentleman, at the University of Auckland, New Zealand, and partly as a play on the name of S language.

The project was conceived in 1992, with an initial version released in 1994 and a stable beta version in 2000.

- easy to install
- easy to learn, very flexible
- friendly graphical user interface (GUI)
- many contributed packages
- extremely popular
- It is free!

Website for downloading R:

https://mirrors.tuna.tsinghua.edu.cn/CRAN/

There are different versions of R for different platforms, for example, windows/(Mac) OS X/linux.

CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage
The R Journal

Software
R Sources
R Binaries
Packages
Other

Documentation
Manuals
FAQs
Contributed

The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- Download R for Linux
- Download R for (Mac) OS X
- Download R for Windows

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

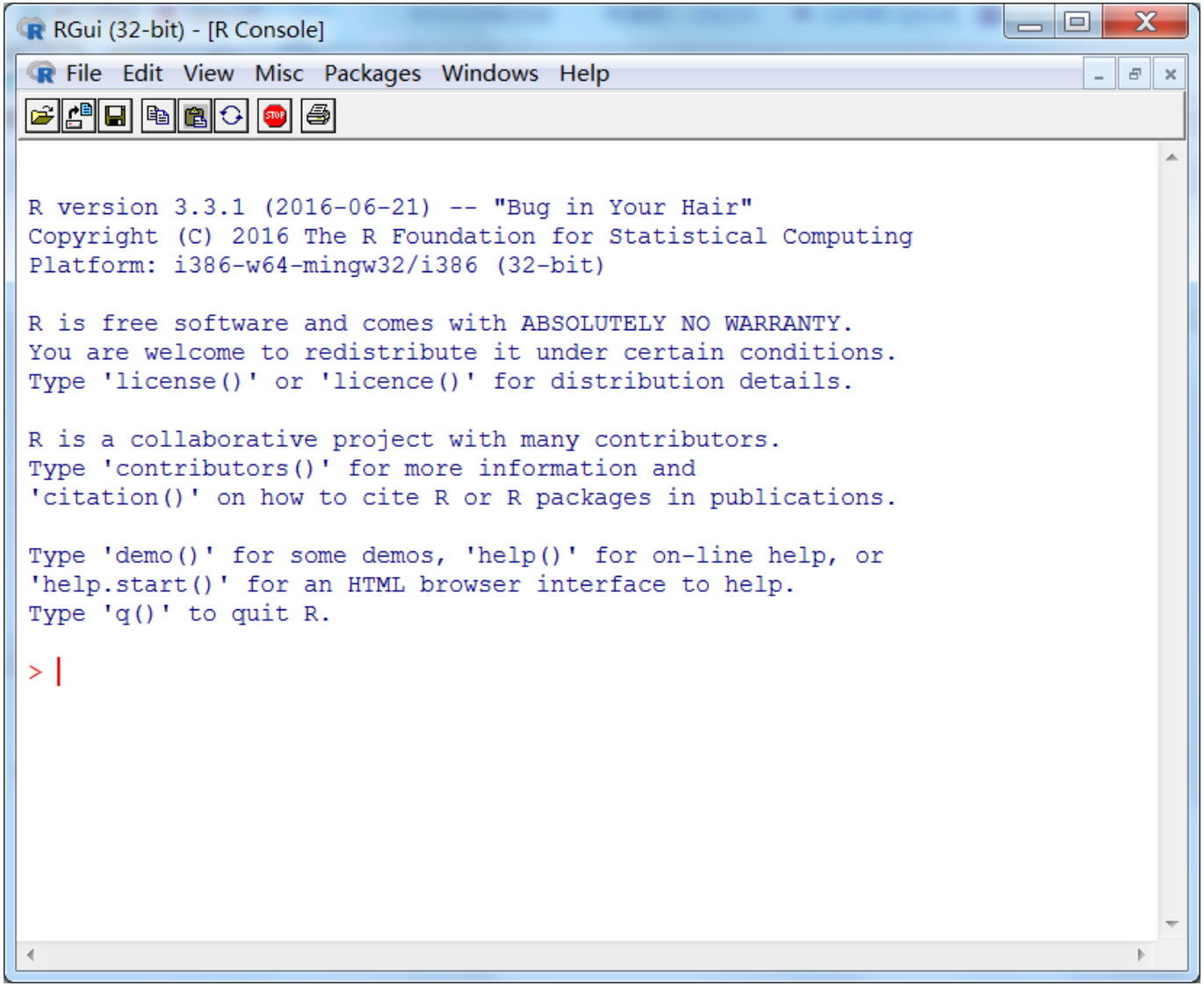
- The latest release (Tuesday 2016-06-21, Bug in Your Hair) R-3.3.1.tar.gz, read what's new in the latest version.
- Sources of R alpha and beta releases (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are available here. Please read about new features and bug fixes before filing corresponding feature requests or bug reports.
- Source code of older versions of R is available here.
- Contributed extension packages

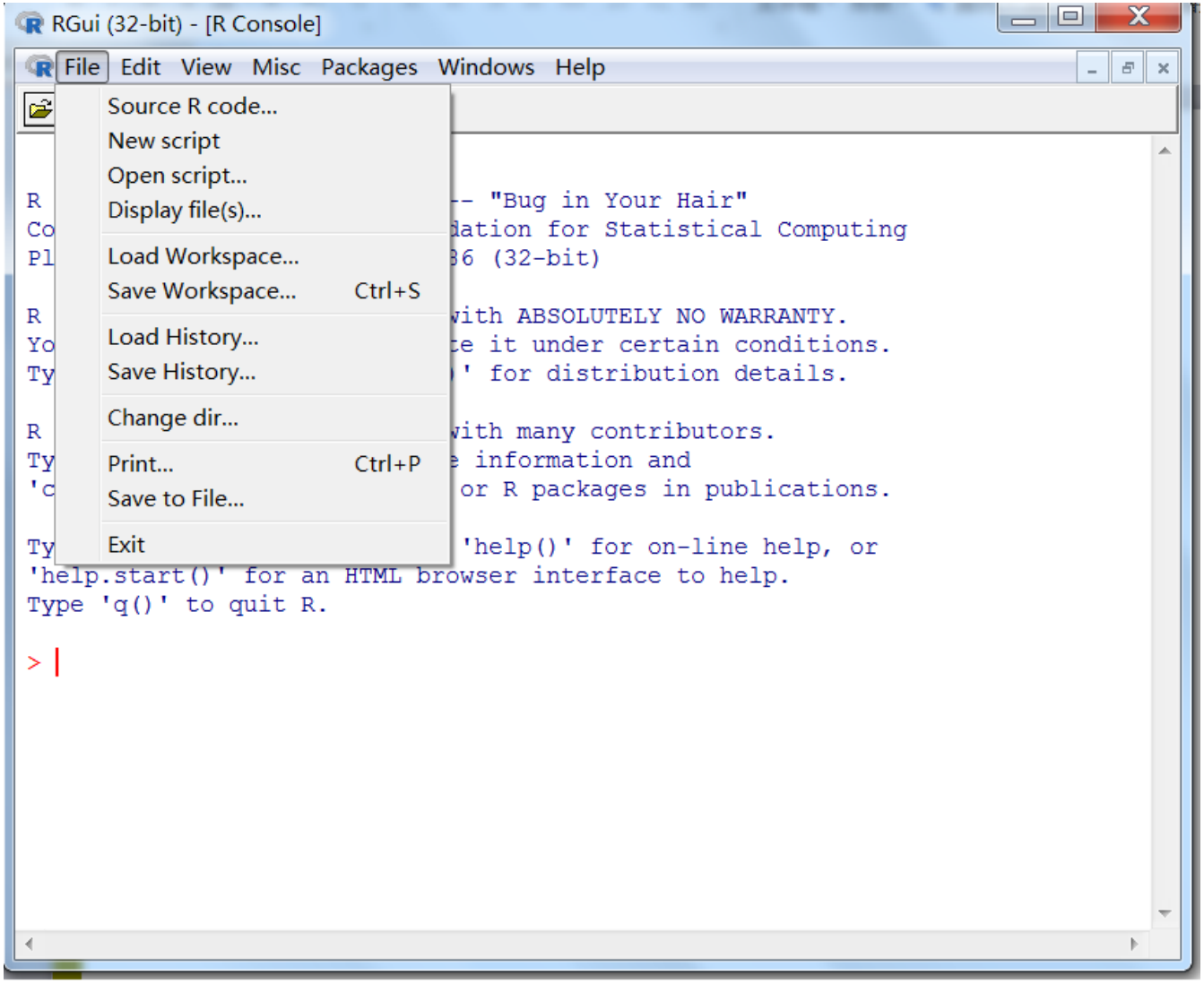
Questions About R

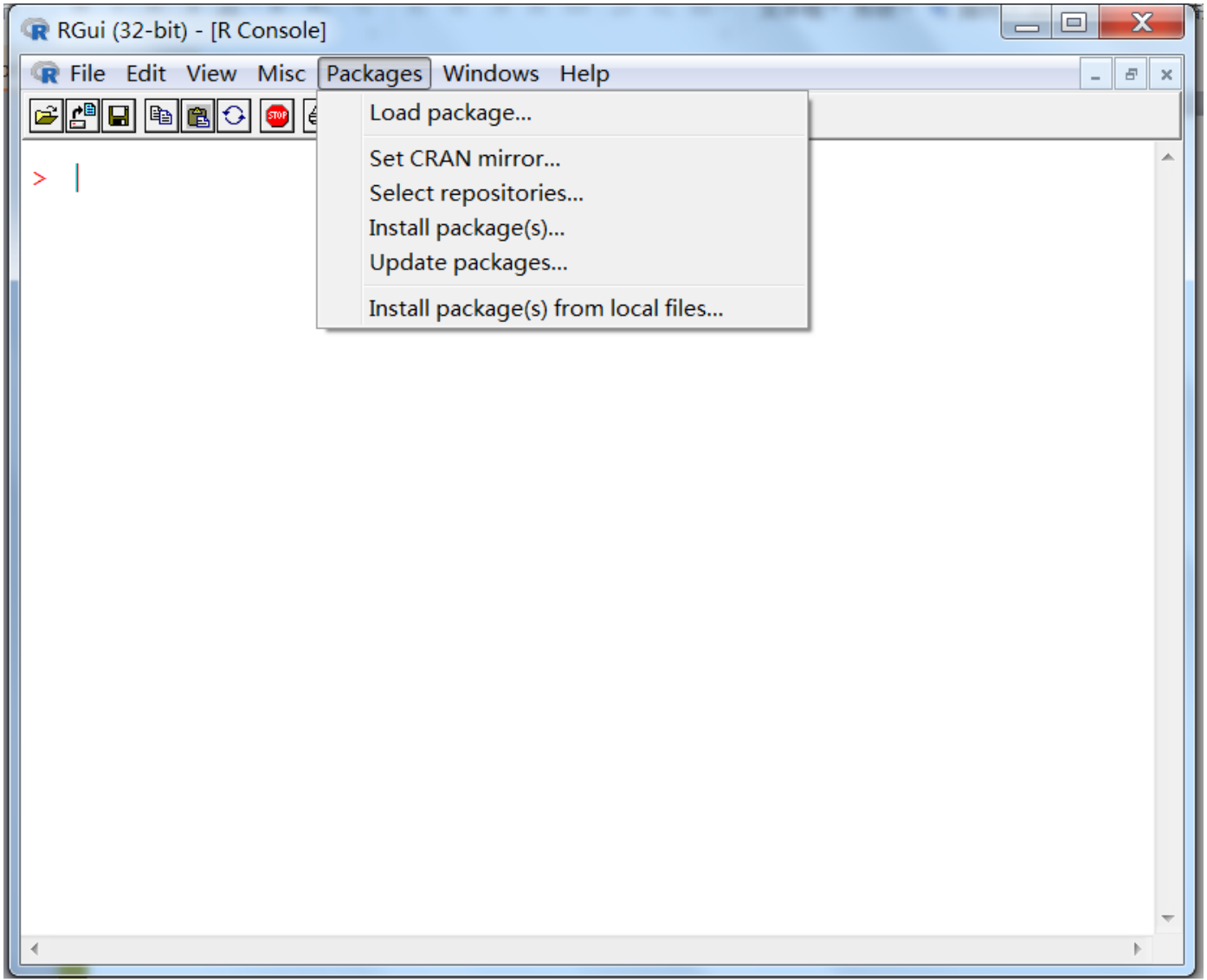
- If you have questions about R like how to download and install the software, or what the license terms are, please read our answers to frequently asked questions before you send an email.

What are R and CRAN?


R i386 3.3.1

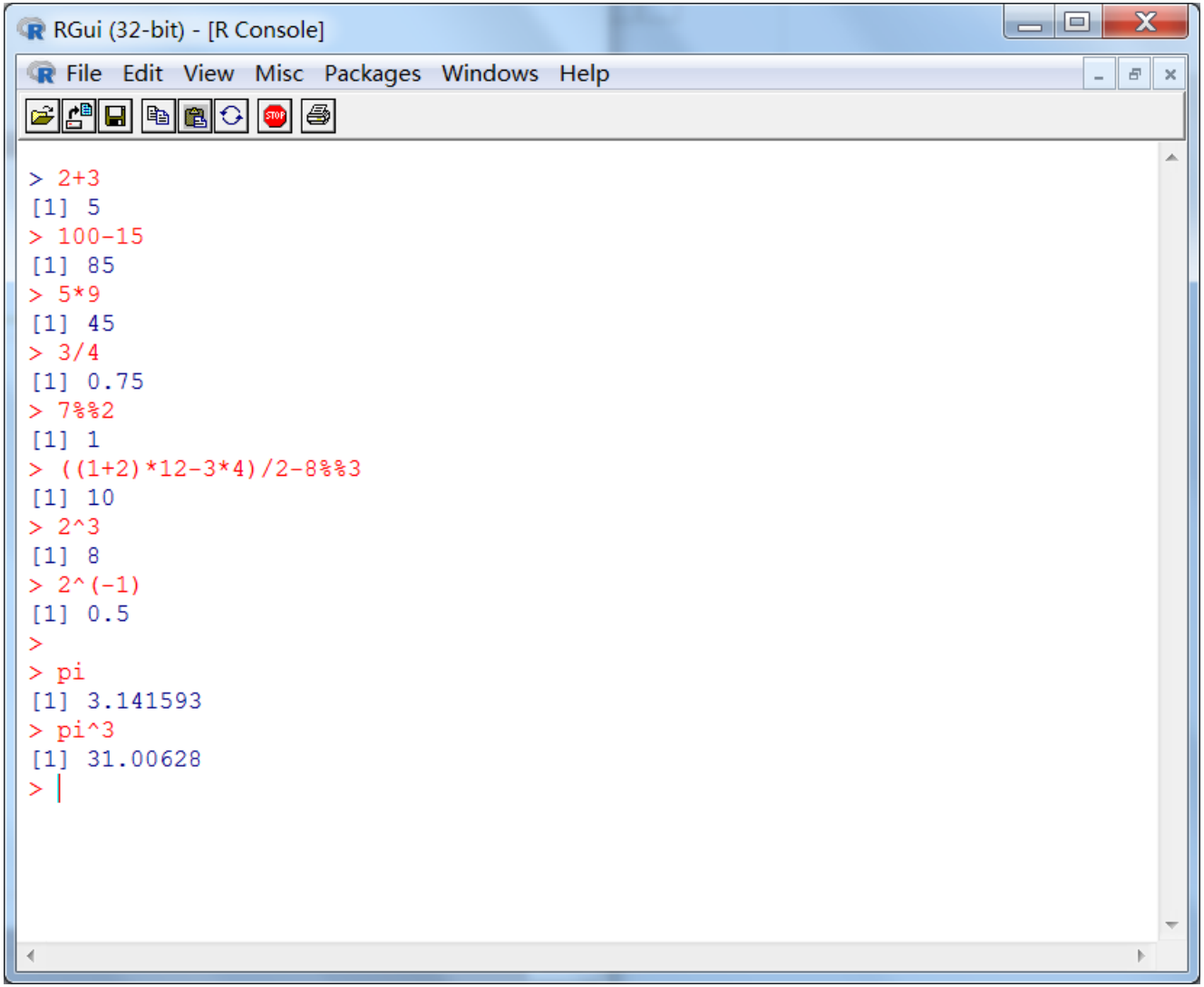

RGui (32-bit) - [R Console]
File Edit View Misc Packages Windows Help
R version 3.3.1 (2016-06-21) -- "Bug in Your Hair"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: i386-w64-mingw32/i386 (32-bit)
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
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.
>


RGui (32-bit) - [R Console]
File Edit View Misc Packages Windows Help
Source R code...
New script
Open script...
Display file(s)...
Load Workspace...
Save Workspace... Ctrl+S
Load History...
Save History...
Change dir...
Print... Ctrl+P
Save to File...
Exit
-- "Bug in Your Hair"
dation for Statistical Computing
86 (32-bit)
with ABSOLUTELY NO WARRANTY.
te it under certain conditions.
' for distribution details.
with many contributors.
e information and
or R packages in publications.
'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
>

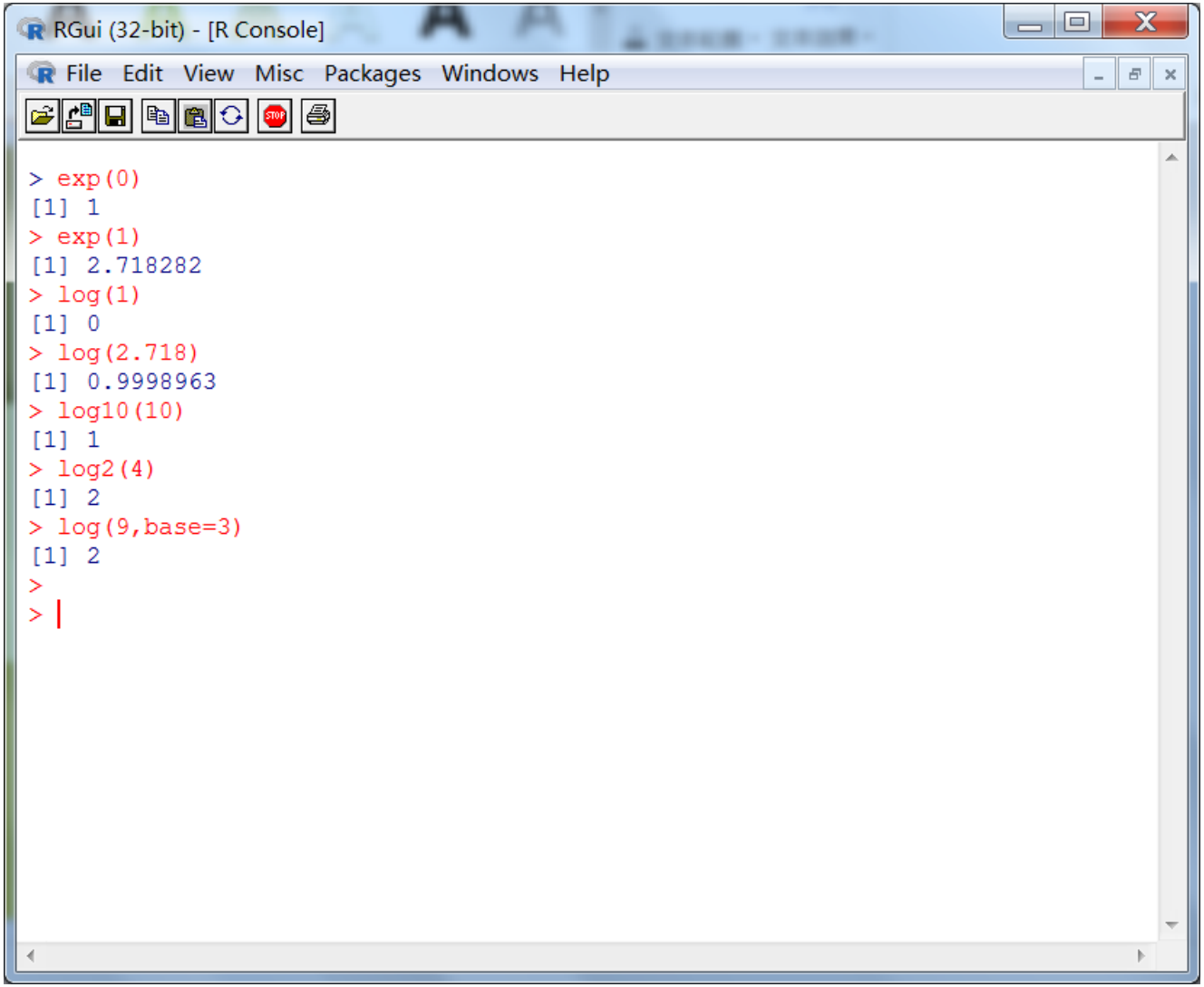

RGui (32-bit) - [R Console]
File Edit View Misc Packages Windows Help
Load package...
Set CRAN mirror...
Select repositories...
Install package(s)...
Update packages...
Install package(s) from local files...
>

```
2+3
100-15
5*9
3/4
7%%2
((1+2)*12-3*4)/2-8%%3
2^3
2^(-1)

pi
pi^3
```


RGui (32-bit) - [R Console]
File Edit View Misc Packages Windows Help
> 2+3
[1] 5
> 100-15
[1] 85
> 5*9
[1] 45
> 3/4
[1] 0.75
> 7%%2
[1] 1
> ((1+2)*12-3*4)/2-8%%3
[1] 10
> 2^3
[1] 8
> 2^(-1)
[1] 0.5
>
> pi
[1] 3.141593
> pi^3
[1] 31.00628
>

```
exp(0)
exp(1)
log(1)
log(2.718)
log10(10)
log2(4)
log(9,base=3)
```


RGui (32-bit) - [R Console]
File Edit View Misc Packages Windows Help
> exp(0)
[1] 1
> exp(1)
[1] 2.718282
> log(1)
[1] 0
> log(2.718)
[1] 0.9998963
> log10(10)
[1] 1
> log2(4)
[1] 2
> log(9,base=3)
[1] 2
>
>

sin(0)

sin(pi/2)

sin(pi)

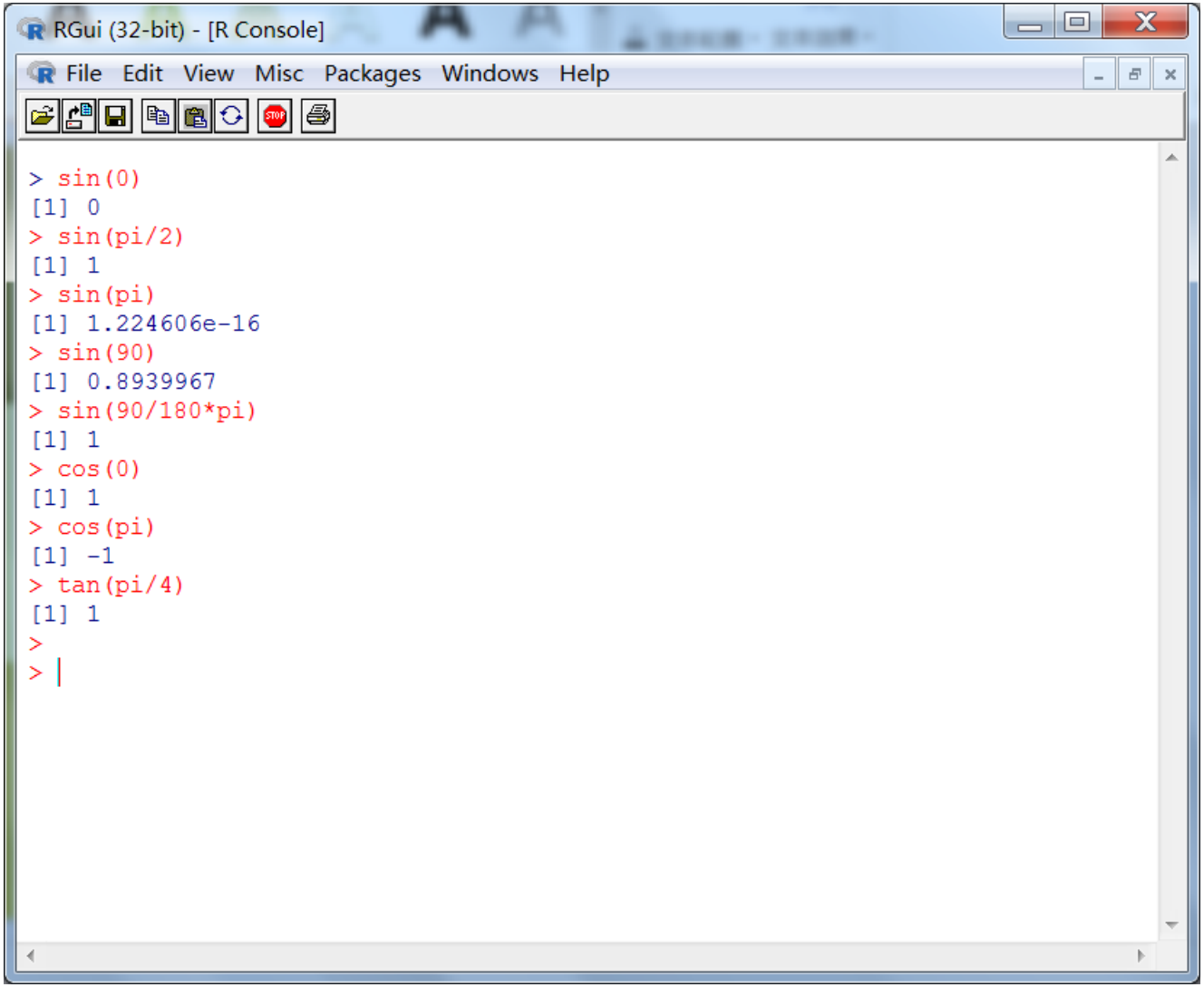
sin(90)

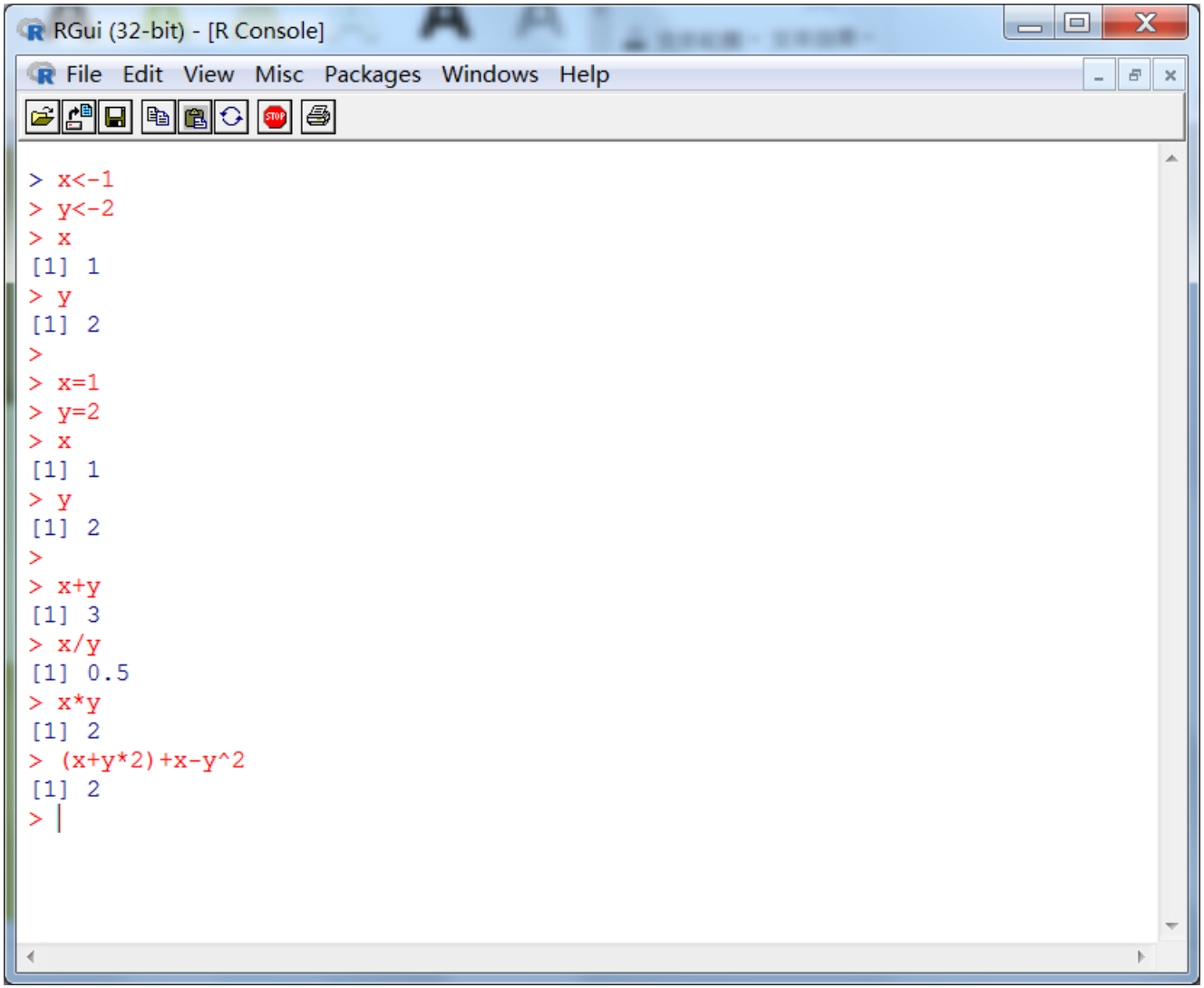
sin(90/180*pi)

cos(0)

cos(pi)

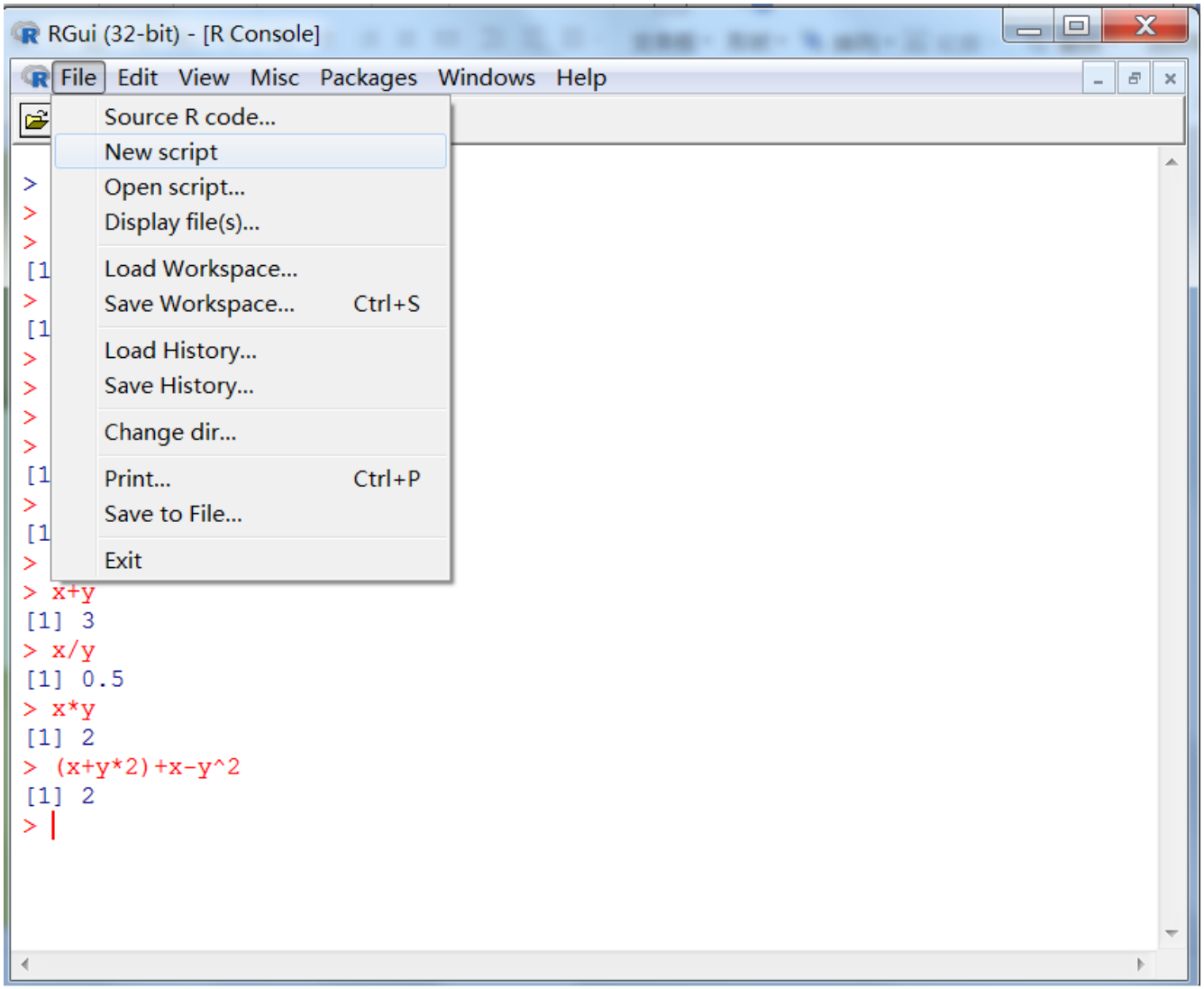
tan(pi/4)

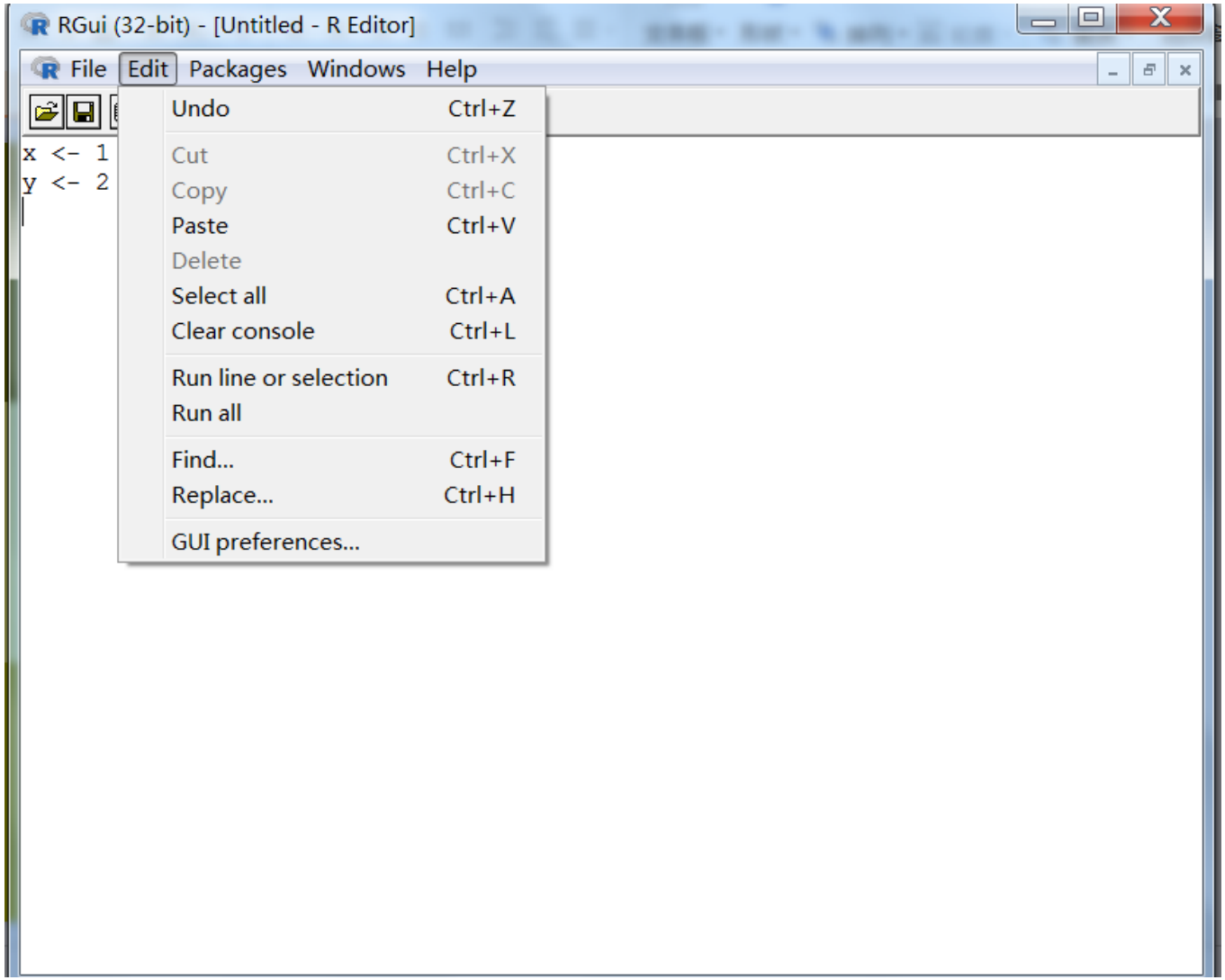

RGui (32-bit) - [R Console]
File Edit View Misc Packages Windows Help
> sin(0)
[1] 0
> sin(pi/2)
[1] 1
> sin(pi)
[1] 1.224606e-16
> sin(90)
[1] 0.8939967
> sin(90/180*pi)
[1] 1
> cos(0)
[1] 1
> cos(pi)
[1] -1
> tan(pi/4)
[1] 1
>
>

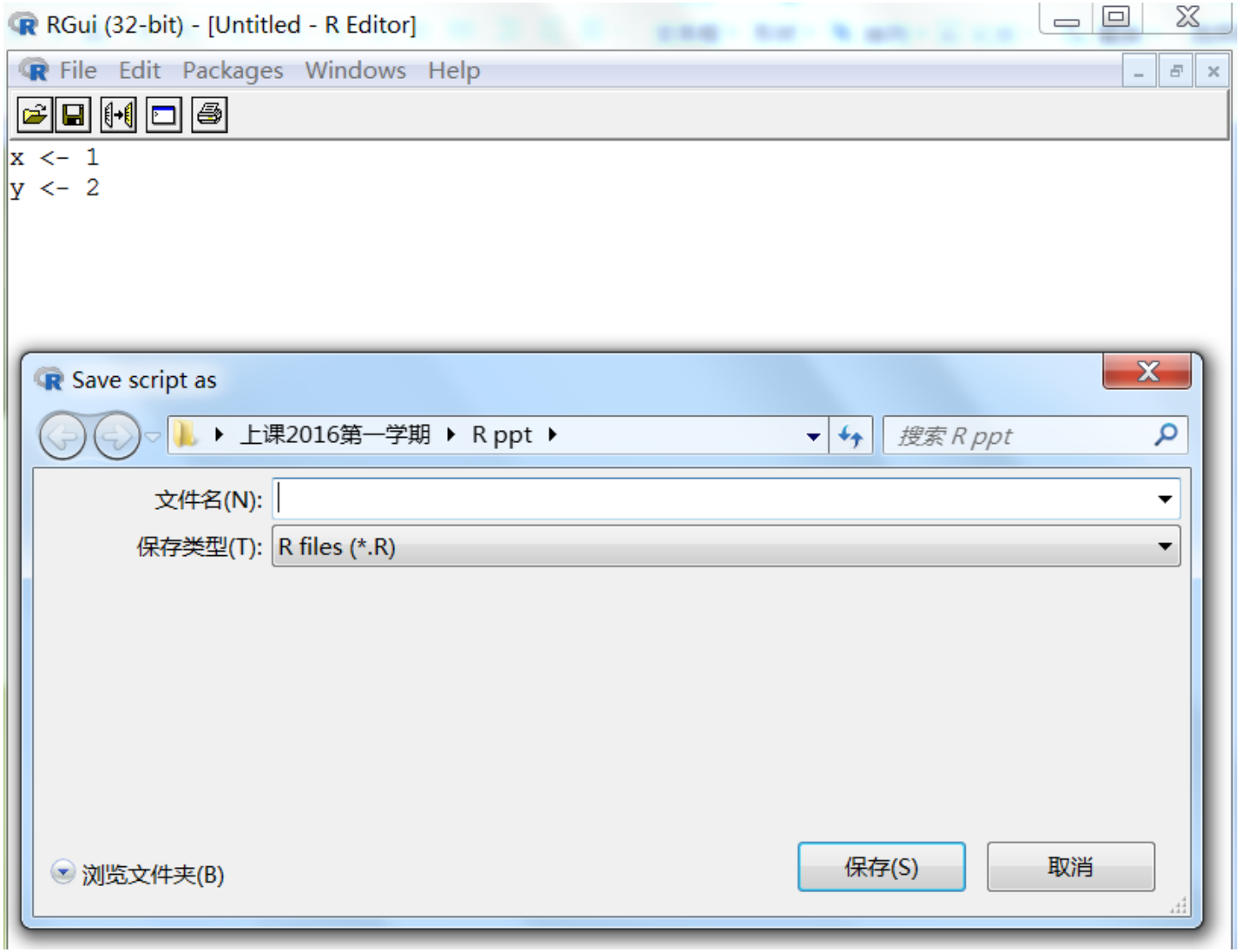

RGui (32-bit) - [R Console]
File Edit View Misc Packages Windows Help
> x<-1
> y<-2
> x
[1] 1
> y
[1] 2
>
> x=1
> y=2
> x
[1] 1
> y
[1] 2
>
> x+y
[1] 3
> x/y
[1] 0.5
> x*y
[1] 2
> (x+y*2)+x-y^2
[1] 2
>

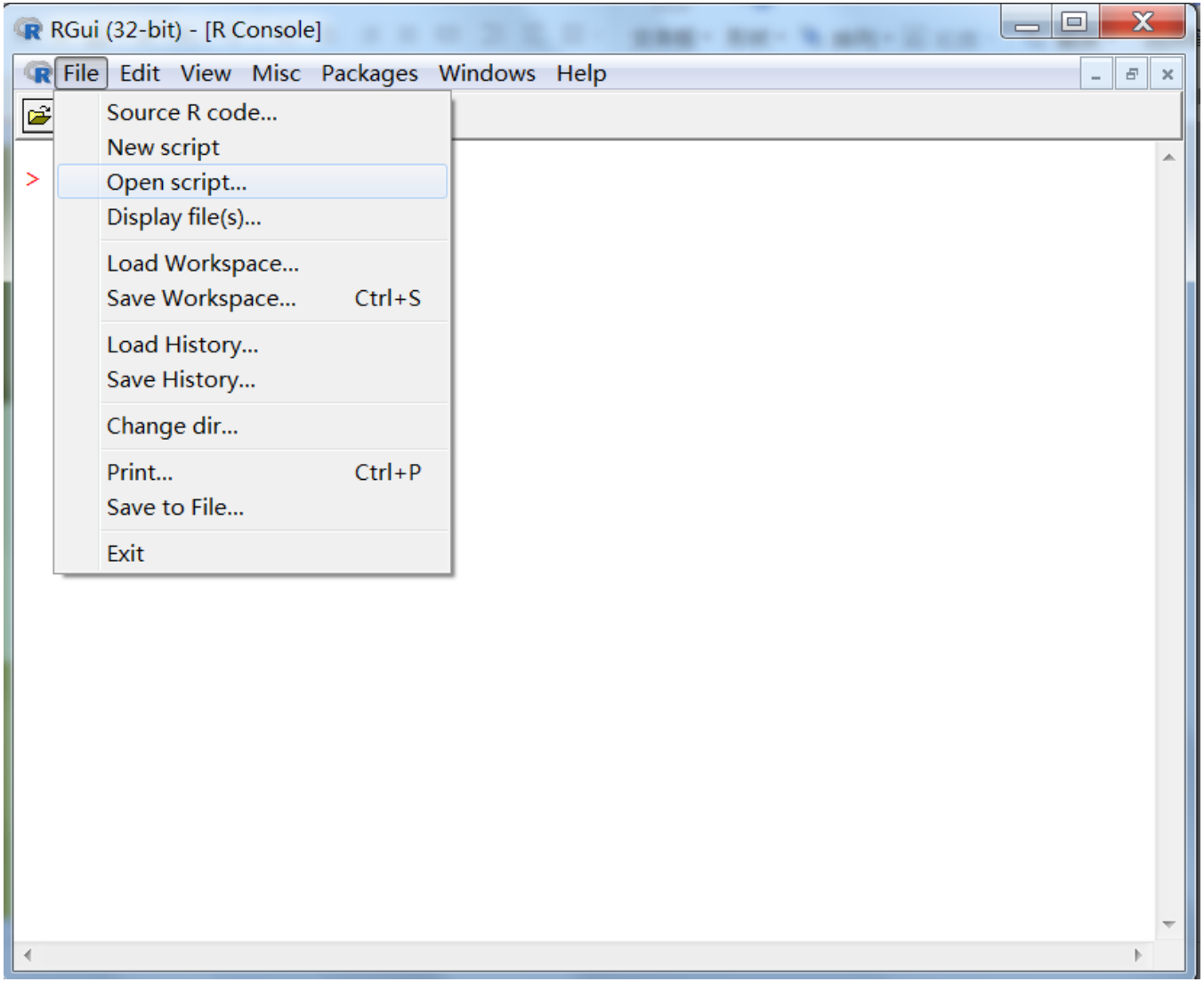
Script- a file ends with .R

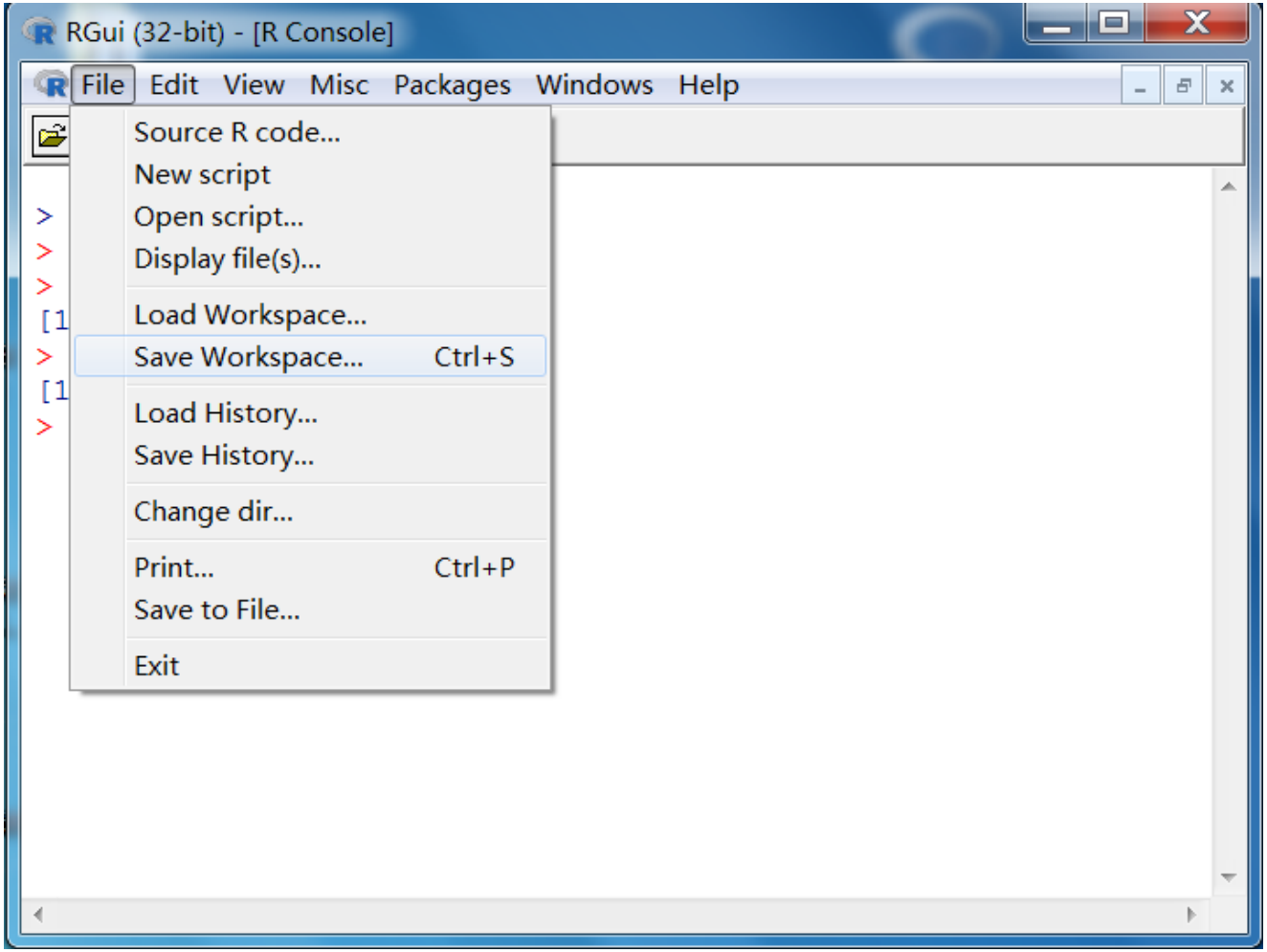
- A script is a file that saves the codes to run.
- It could be a few lines, or a very large text file.
- It could be edited like a text file, for example, copy, paste, delete ....
- When running the R script, the console will run the codes line by line.


RGui (32-bit) - [R Console]
File Edit View Misc Packages Windows Help
Source R code...
New script
Open script...
Display file(s)...
Load Workspace...
Save Workspace... Ctrl+S
Load History...
Save History...
Change dir...
Print... Ctrl+P
Save to File...
Exit
> x+y
[1] 3
> x/y
[1] 0.5
> x*y
[1] 2
> (x+y*2)+x-y^2
[1] 2
>


RGui (32-bit) - [Untitled - R Editor]
File
Edit
Packages
Windows
Help
x <- 1
y <- 2
Undo Ctrl+Z
Cut Ctrl+X
Copy Ctrl+C
Paste Ctrl+V
Delete
Select all Ctrl+A
Clear console Ctrl+L
Run line or selection Ctrl+R
Run all
Find... Ctrl+F
Replace... Ctrl+H
GUI preferences...

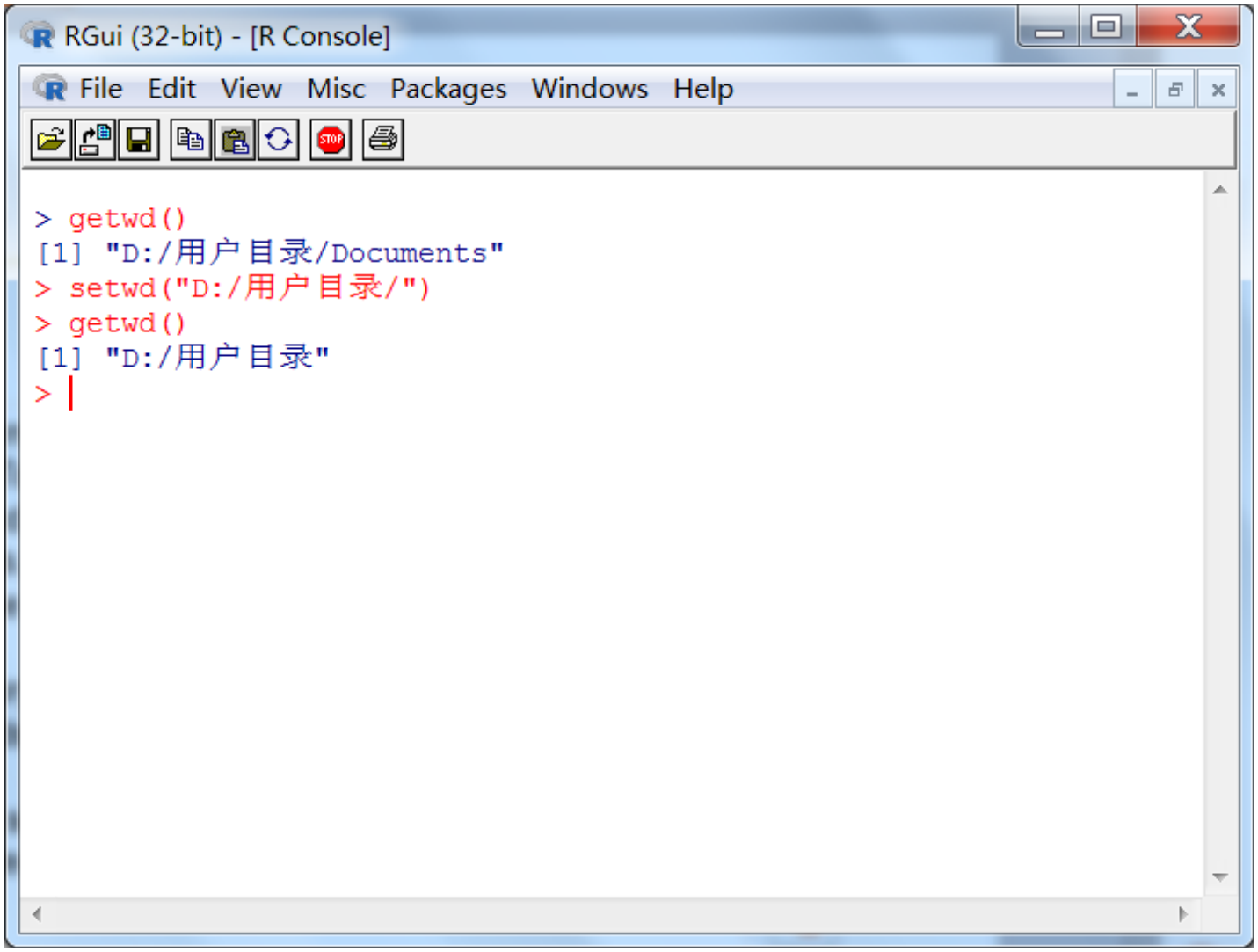

RGui (32-bit) - [Untitled - R Editor]
File Edit Packages Windows Help
x <- 1
y <- 2
Save script as
上课2016第一学期 ▸ R ppt ▸
搜索 R ppt
文件名(N):
保存类型(T):
R files (*.R)
浏览文件夹(B)
保存(S)
取消


RGui (32-bit) - [R Console]
File
Edit
View
Misc
Packages
Windows
Help
Source R code...
New script
Open script...
Display file(s)...
Load Workspace...
Save Workspace... Ctrl+S
Load History...
Save History...
Change dir...
Print... Ctrl+P
Save to File...
Exit
>


RGui (32-bit) - [R Console]
File
Edit
View
Misc
Packages
Windows
Help
Source R code...
New script
Open script...
Display file(s)...
Load Workspace...
Save Workspace... Ctrl+S
Load History...
Save History...
Change dir...
Print... Ctrl+P
Save to File...
Exit

Working directory is the current environment where R works, saves and reads varaibles or data by default.

Use getwd() to see the working directory, setwd() to change the working directory.


RGui (32-bit) - [R Console]
File Edit View Misc Packages Windows Help
> getwd()
[1] "D:/用户目录/Documents"
> setwd("D:/用户目录/")
> getwd()
[1] "D:/用户目录"
>

History

- history() will show all the codes that were run in the console.
- Click “File - Save History”, then save the codes that were run in a file end with “.RHistory”.
- Click “File - Load History”, then load a R history file.

We can also save the history codes as a text file.

- Use “history()” to see all the codes in the R history window.
- Under the R history window, click “File - Save to File”, then save the codes as a file ended with “.txt”.

Package

R comes with extensive capabilities right out of optional modules that you can download and install. There are over 2,500 user-contributed modules called packages that you can download from

http://cran.r-project.org/web/packages

To see the packages already installed in R, use

- library() or installed.packages()

To see the packages already loaded in R, use

- search()

How to install a package

How to install a package called “gclus”? There are two ways.

- Use the code directly as follows.

  install.packages(“gclus”)

- Manually by clicking “Packages - install package(s)”, then choose the package “gclus” to install.

How to load a packge

After installing the “gclus” package, if you want to use it, you need to use

- library(“gclus”)

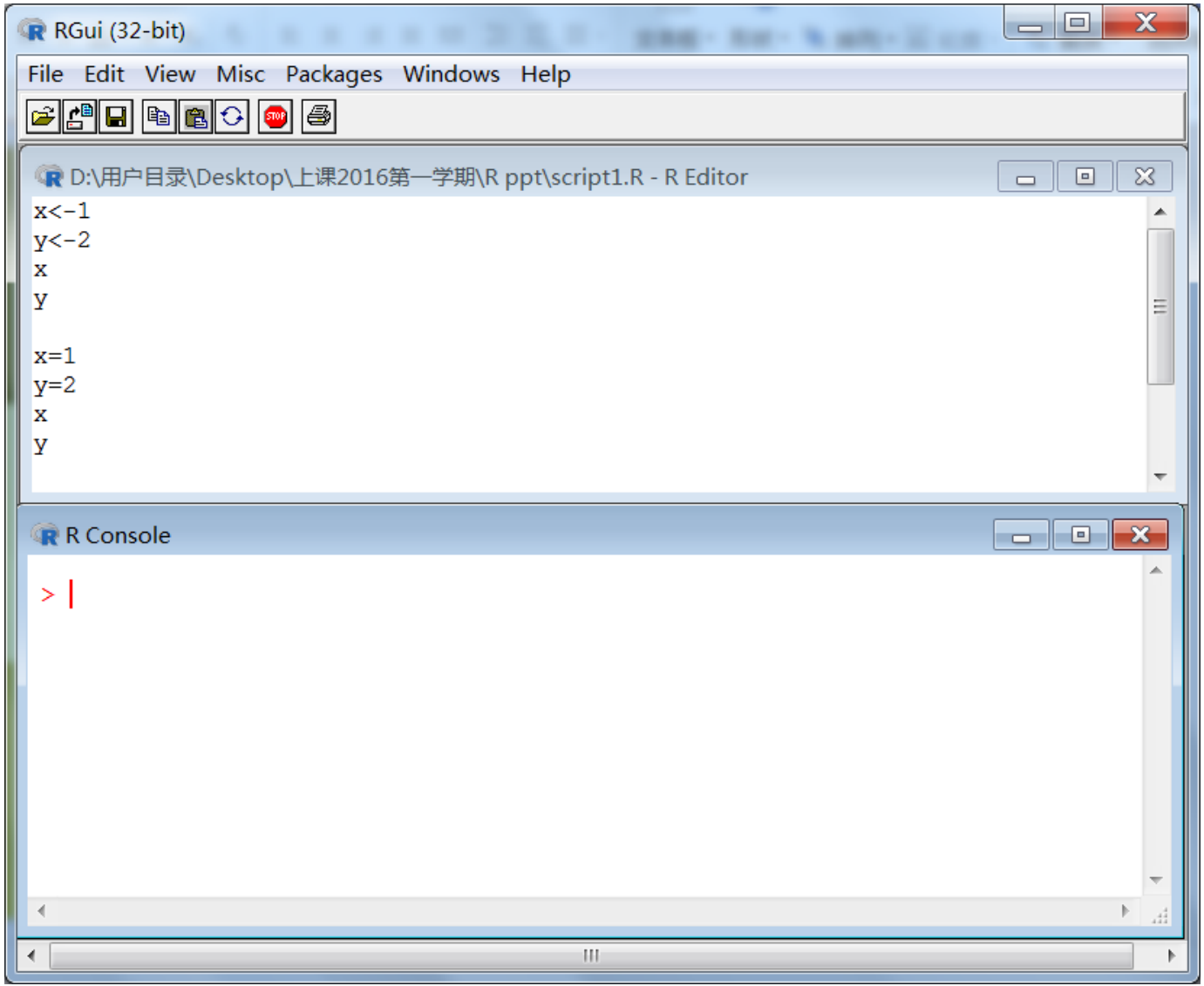
Then package “gclus” is loaded in the current environment, and you can acess its data and functions now.

help()

If you want to see the details of a R function, for example, “mean”, use

- help(mean) or ?mean

- Ctrl+L= clear the console
- Ctrl+C=copy
- Ctrl+V=paste
- ls()=show all the existing variables
- rm(x)= delete the variable x
- rm(x,y)= delete the variable x and y


RGui (32-bit)
File Edit View Misc Packages Windows Help
D:\用户目录\Desktop\上课2016第一学期\R ppt\script1.R - R Editor
x<-1
y<-2
x
y
x=1
y=2
x
y
R Console
>

- how to install R
- R's advantages
- R's GUI environment
- Use R to do calculation
- build, open and save R's script/workspace/history
- install/load a R package
- use help()
- some tips

Begin the journey to explore R. Have fun!