

강원대학교  
AI 소프트웨어학과

---

# 머신러닝1

---

## 강의목표

- **통계를 사용하는 이유와 학습 방법에 대한 이해**
- **통계에 대한 이해를 통해 문제 해결**
- **실제 데이터를 실습하며 방법론에 대한 이해**
- **실제로 데이터가 만들어지고 학습되는 과정을 이해**
- **소규모 프로젝트를 통해 새로운 문제를 해결**

- 성적은 중간고사 : 시험, 기말고사 : 시험
- 지각 3번은 결석 1번
- 상대평가
- 중간고사 또는 기말고사 미응시 : F
- 결석 4번은 F



Search or jump to...



[Pull requests](#)

[Issues](#)

[Codespaces](#)

[Marketplace](#)

[Explore](#)



[Overview](#)

[Repositories](#) 7

[Projects](#)

[Packages](#)

[Stars](#) 1



Kimchanggyun

[Edit profile](#)

0 followers · 1 following

**Organizations**



### Popular repositories

[Customize your pins](#)

[MLPB](#)

Forked from ben519/MLPB

Machine Learning Problem Bible | Problem Set Here >>



R

Public

[facerecognition](#)

Forked from bwcho75/facerecognition



Jupyter Notebook

Public

[Tensorflow-101](#)

Forked from sjchoi86/Tensorflow-101

TensorFlow Tutorials



Jupyter Notebook

Public

[models](#)

Forked from tensorflow/models

Models and examples built with TensorFlow



Python

Public

[soynlp](#)

Forked from lovit/soynlp

한국어 자연어처리를 위한 파이썬 라이브러리입니다. 단어 추출/ 토큰나이저 / 품사판별/ 전처리의 기능을 제공합니다.



Jupyter Notebook

Public

3 contributions in the last year

[Contribution settings](#)

Overview Repositories 7 Projects Packages Stars 1

### Popular repositories

MLPB

Forked from ben519/MLPB

Machine Learning Problem Bible | Problem Set Here >>

R

Public

facerecognition

Forked from bwcho75/facerecognition

Jupyter Notebook

Public

Customize your pins

Kimchanggyun

Type to search

Overview Repositories 8 Projects Packages Stars 1

Popular repositories

MLPB

Forked from ben519/MLPB

Machine Learning Problem Bible | Problem Set Here >>

R

Public

facerecognition

Forked from bwcho75/facerecognition

Jupyter Notebook

Public

Customize your pins

- New repository
- Import repository
- New codespace
- New gist
- New organization
- New project

3 contributions in the last year


Contribution settings

# Create a new repository

A repository contains all project files, including the revision history. Already have a project repository elsewhere? [Import a repository.](#)

Required fields are marked with an asterisk (\*).

Owner \*

 Kimchanggyun

 / 

Repository name \*


R

✔ R is available.

Great repository names are short and memorable. Need inspiration? How about [laughing-garbanzo](#) ?

Description (optional)


- ☒



Public

Anyone on the internet can see this repository. You choose who can commit.

☐



Private



You choose who can see and commit to this repository.
- Initialize this repository with:

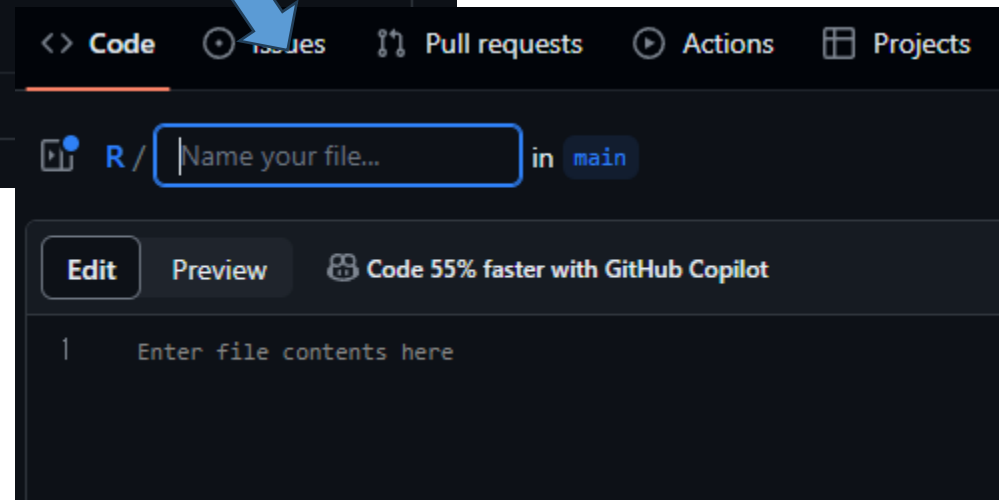
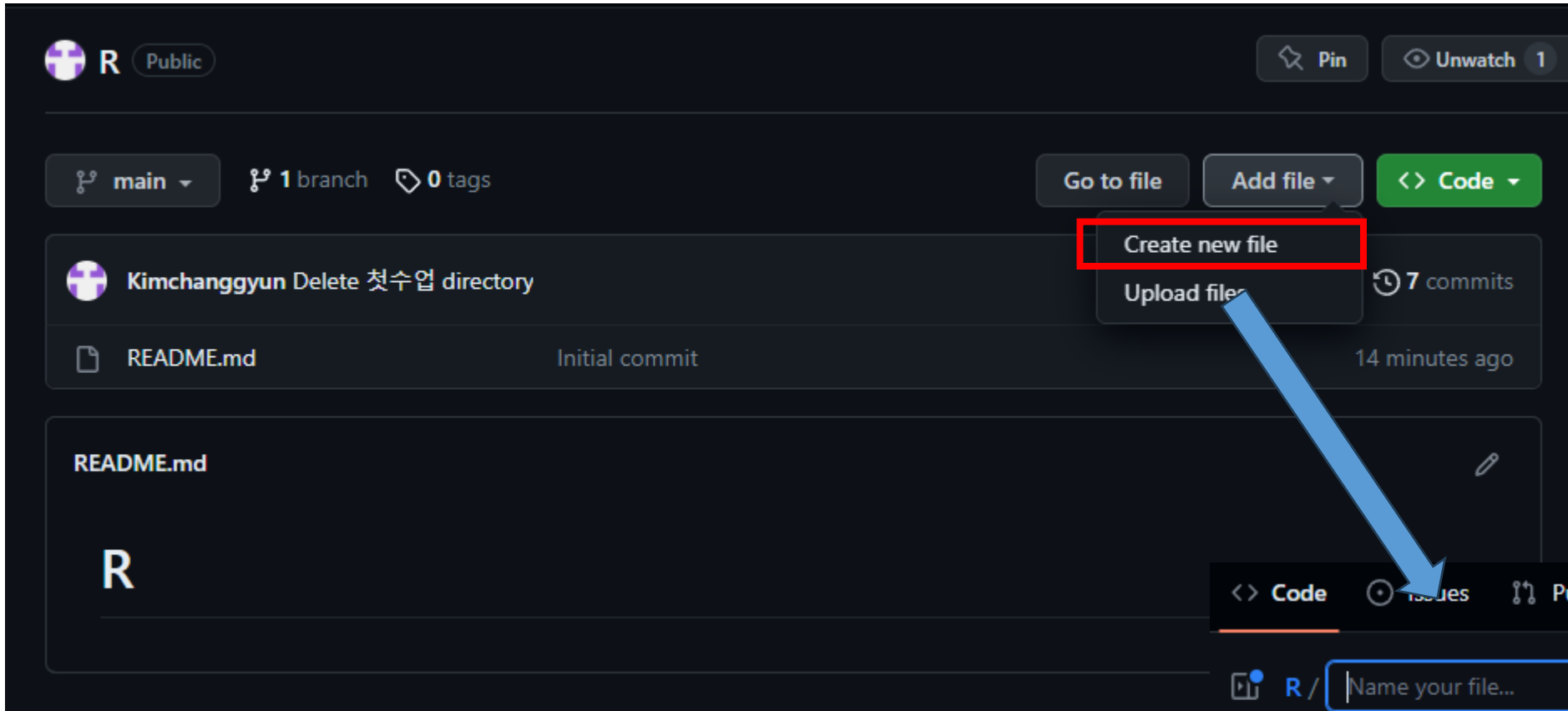
☒

Add a README file

This is where you can write a long description for your project. [Learn more about READMEs.](#)
- Add .gitignore

.gitignore template: None
- Choose which files not to track from a list of templates. [Learn more about ignoring files.](#)
- Choose a license

License: None
- A license tells others what they can and can't do with your code. [Learn more about licenses.](#)
- This will set  main as the default branch. Change the default name in your [settings](#).
-  You are creating a public repository in your personal account.
- Create repository



## Danger Zone

**Change repository visibility**  
This repository is currently public.

Change visibility

**Transfer ownership**  
Transfer this repository to another user or to an organization where you have the ability to create repositories.

Transfer

**Archive this repository**  
Mark this repository as archived and read-only.

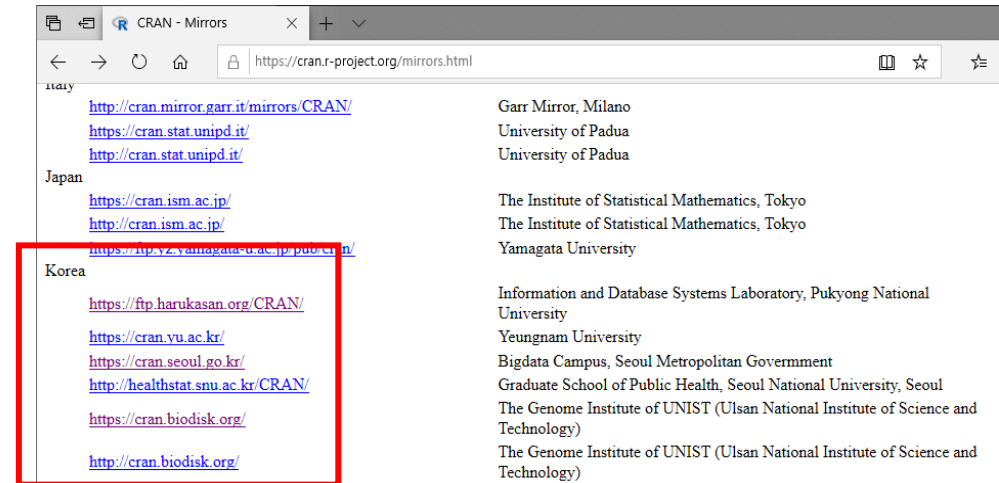
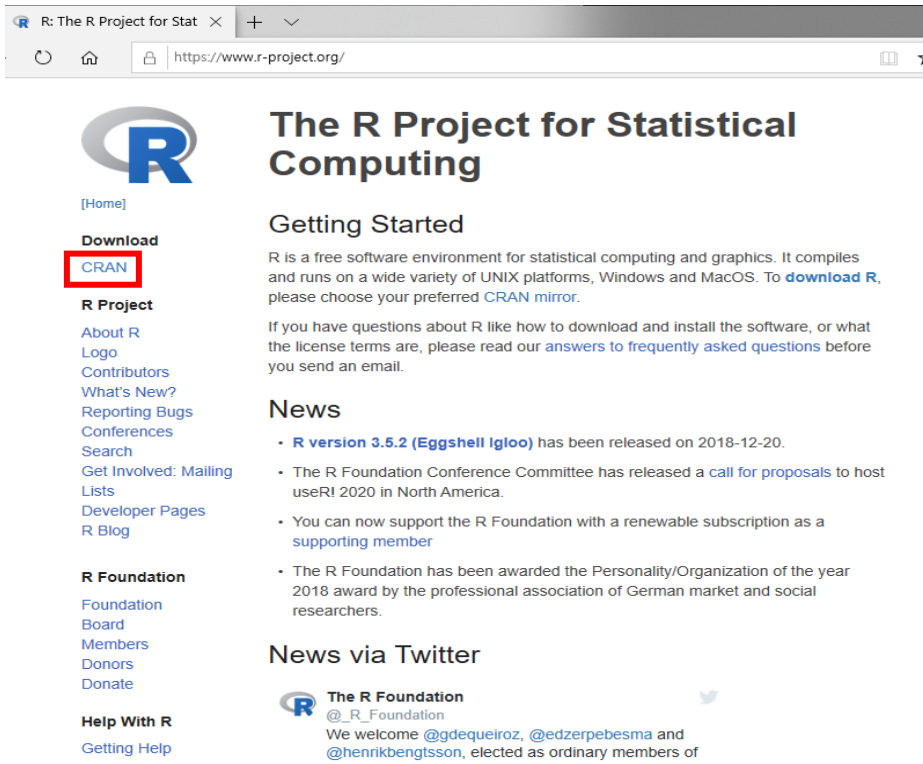
Archive this repository

**Delete this repository**  
Once you delete a repository, there is no going back. Please be certain.

Delete this repository



# R 소개 및 설치

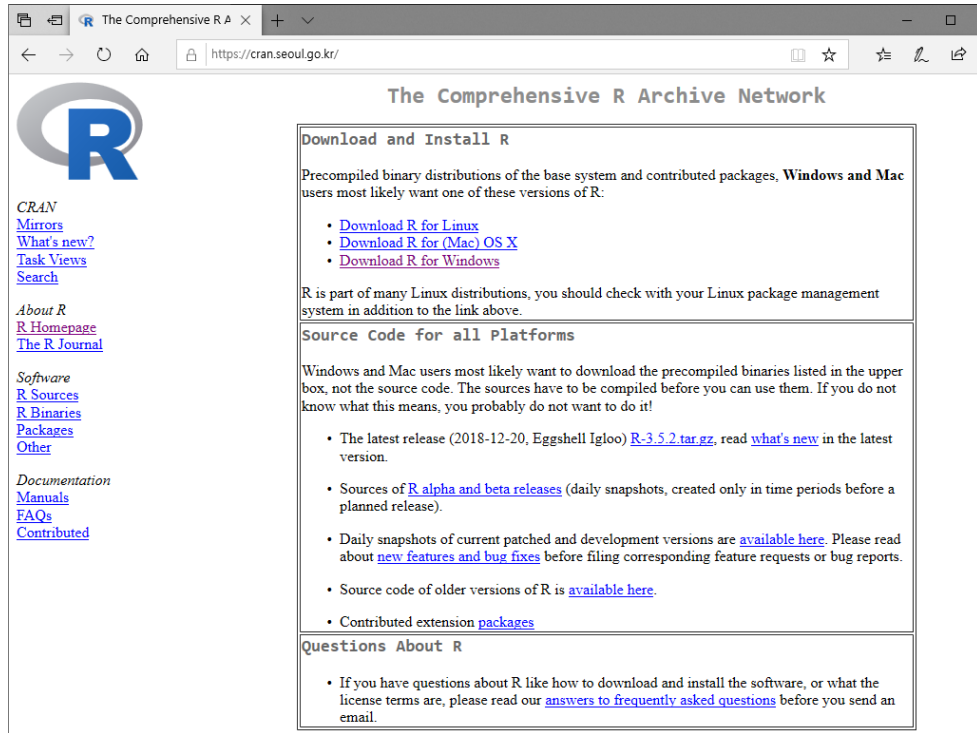


① R 홈페이지 접속 -> <http://www.r-project.org/>

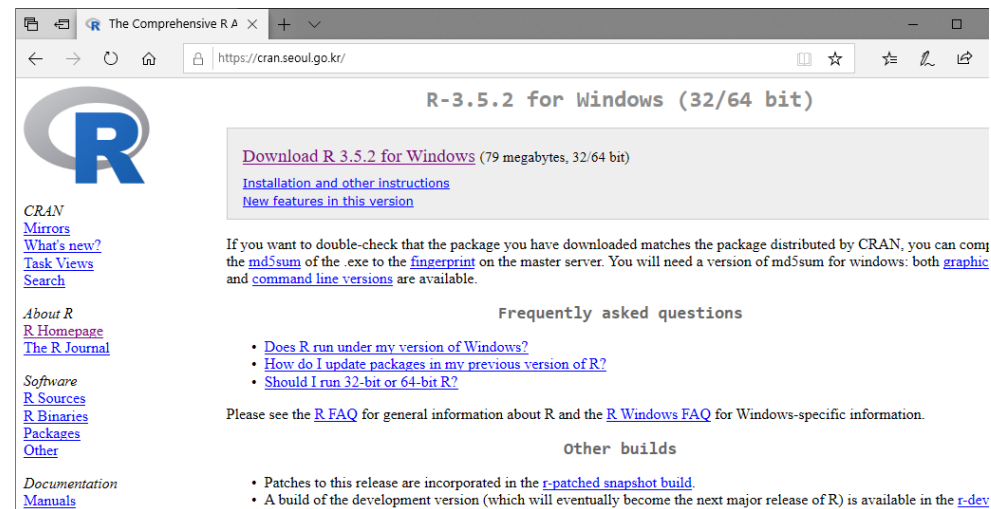
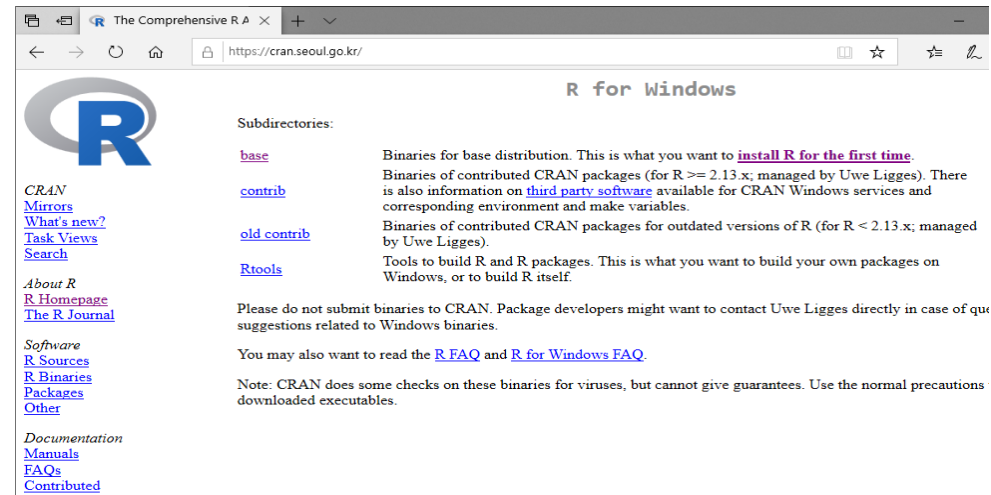
② Download 메뉴의 CRAN 클릭

③ 해당하는 국가(Korea)의 링크 클릭

# R 소개 및 설치

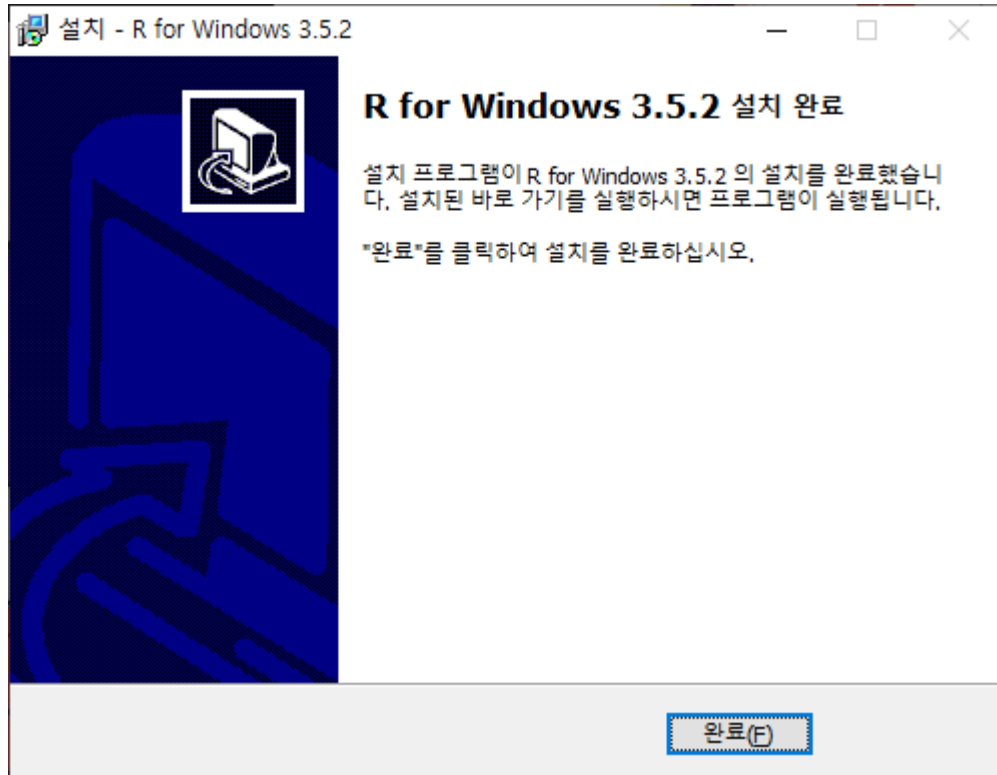


## ④ Download R for Windows 클릭

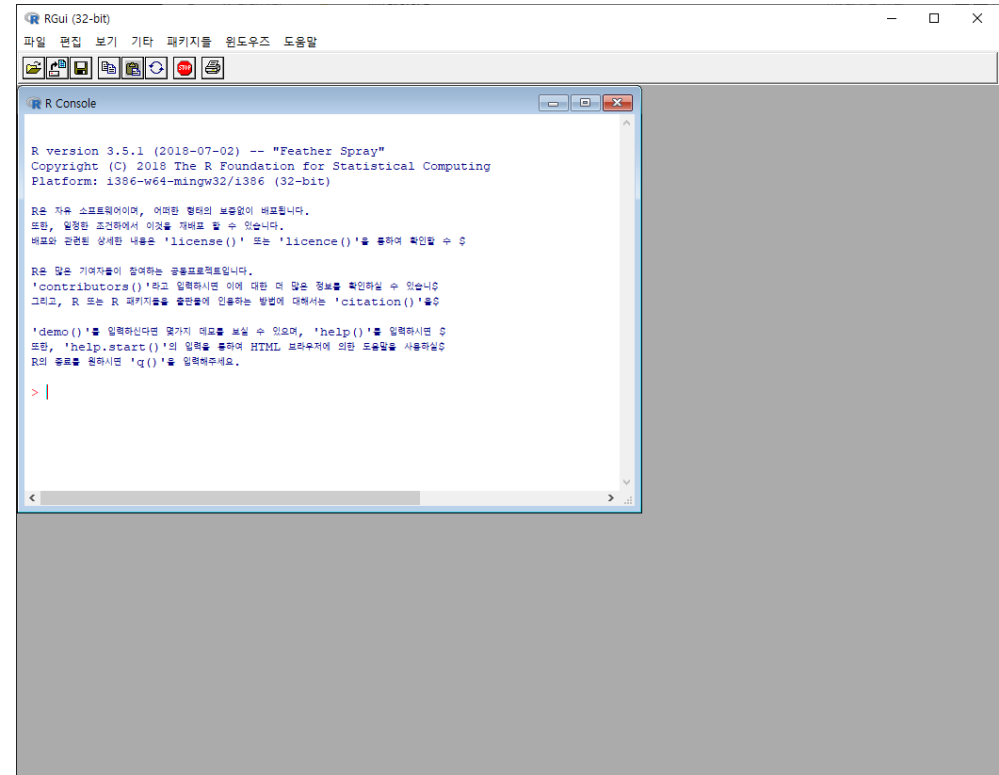


## ⑤ 'base' 클릭

## ⑥ Download R 3.5.2 for Windows를 클릭

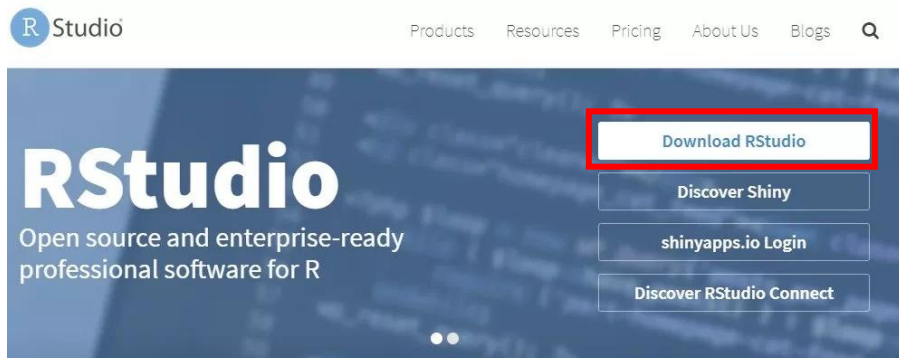


⑦ R 설치완료



⑧ R 실행해보기

# R 소개 및 설치



## RStudio

RStudio makes R easier to use. It includes a code editor, debugging & visualization tools.

[Download](#) [Learn More](#)



## Shiny

Shiny helps you make interactive web applications for visualizing data. Bring R data analysis to life.

[Learn More](#)

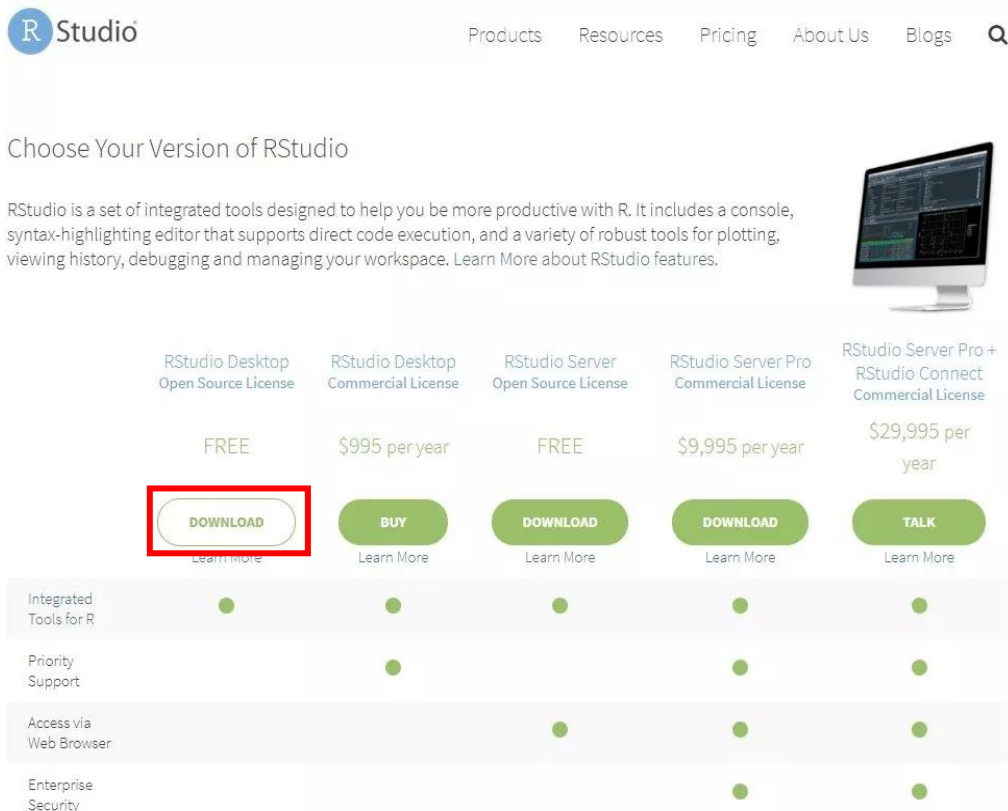


## R Packages

Our developers create popular packages to expand the features of R. Includes ggplot2, dplyr, R Markdown & more.

[Learn More](#)

※ R Studio는 R을 설치하지 않으면 실행 되지 않음

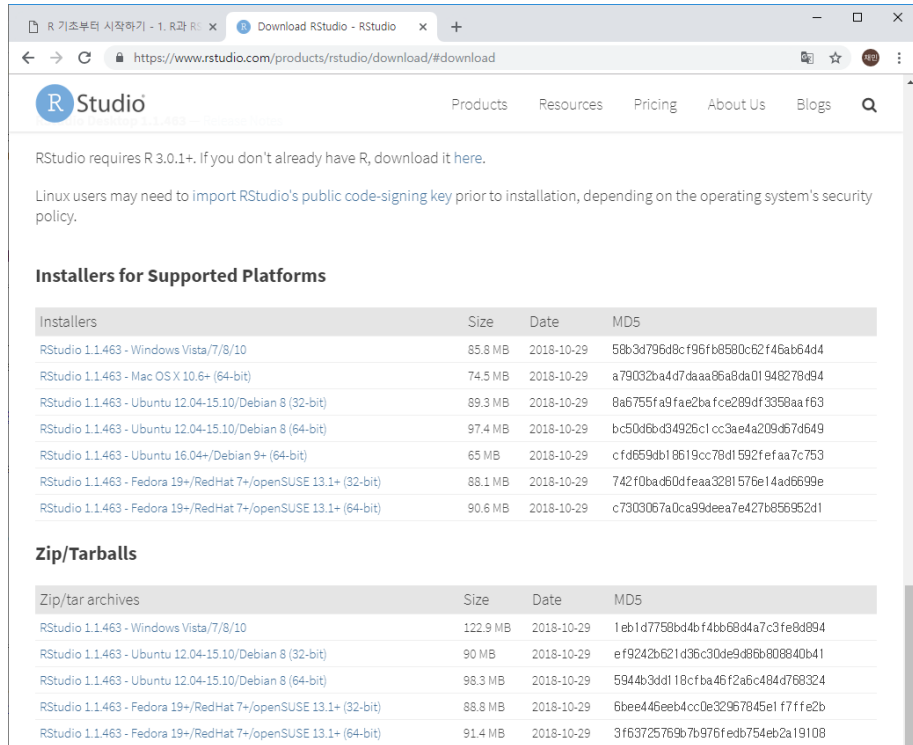


① R Studio 홈페이지 접속 -> <http://www.rstudio.com/>

③ Free 아래의 DOWNLOAD 클릭

② 'Download Rstudio' 클릭

# R 소개 및 설치



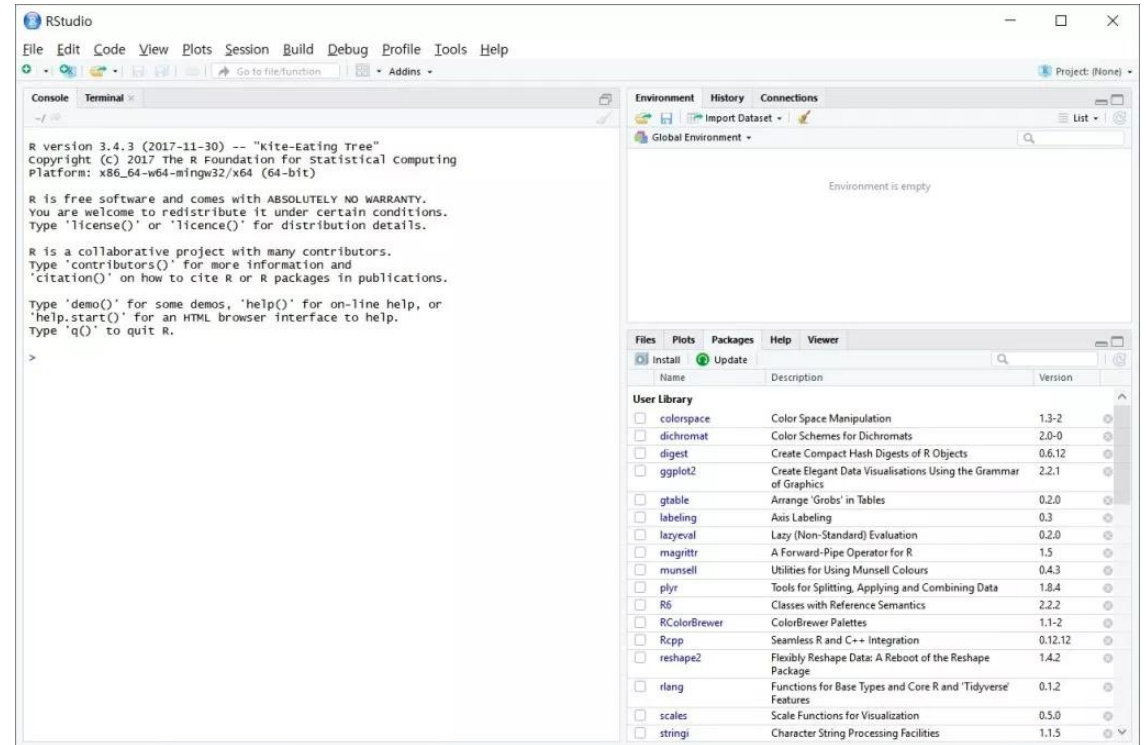
The screenshot shows the RStudio website's download page. It includes navigation links like Products, Resources, Pricing, About Us, and Blogs. A message states: "RStudio requires R 3.0.1+. If you don't already have R, download it here." Below this, a note for Linux users mentions the need to import RStudio's public code-signing key. The main section is titled "Installers for Supported Platforms" and contains a table with download links, sizes, dates, and MD5 hashes for various operating systems. A second table titled "Zip/Tarballs" provides similar information for source code archives.

Installers	Size	Date	MD5
RStudio 1.1.463 - Windows Vista/7/8/10	85.8 MB	2018-10-29	58b3d796d8cf96fb8580c62f46ab6444
RStudio 1.1.463 - Mac OS X 10.6+ (64-bit)	74.5 MB	2018-10-29	a79032ba4d7daa96a8da01948278d94
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	89.3 MB	2018-10-29	0a6755fa9fae2bafce289df3358aa f63
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	97.4 MB	2018-10-29	bc50d6bd34926c1cc3ae4a209d57d649
RStudio 1.1.463 - Ubuntu 16.04+/Debian 9+ (64-bit)	65 MB	2018-10-29	cf6599b18619cc78d1592fefaa7c753
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	88.1 MB	2018-10-29	742f0ba6d0dfeaa3281576e14ad6699e
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	90.6 MB	2018-10-29	c7303067a0ca99deea7e4276956952d1

Zip/tar archives	Size	Date	MD5
RStudio 1.1.463 - Windows Vista/7/8/10	122.9 MB	2018-10-29	1eb1d7758bd4bf4bb68d4a7c3fe8d994
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	90 MB	2018-10-29	ef9242b621d36c30de9d8b808840b41
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	98.3 MB	2018-10-29	5944b3dd118cfba46f2a6c484d768324
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	88.8 MB	2018-10-29	6bee445eeb4cc0e32967845e1f7ffe2b
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	91.4 MB	2018-10-29	3f63725769b7b976fedb754eb2a19108

④ 맨 위에 있는 Windows 버전 클릭

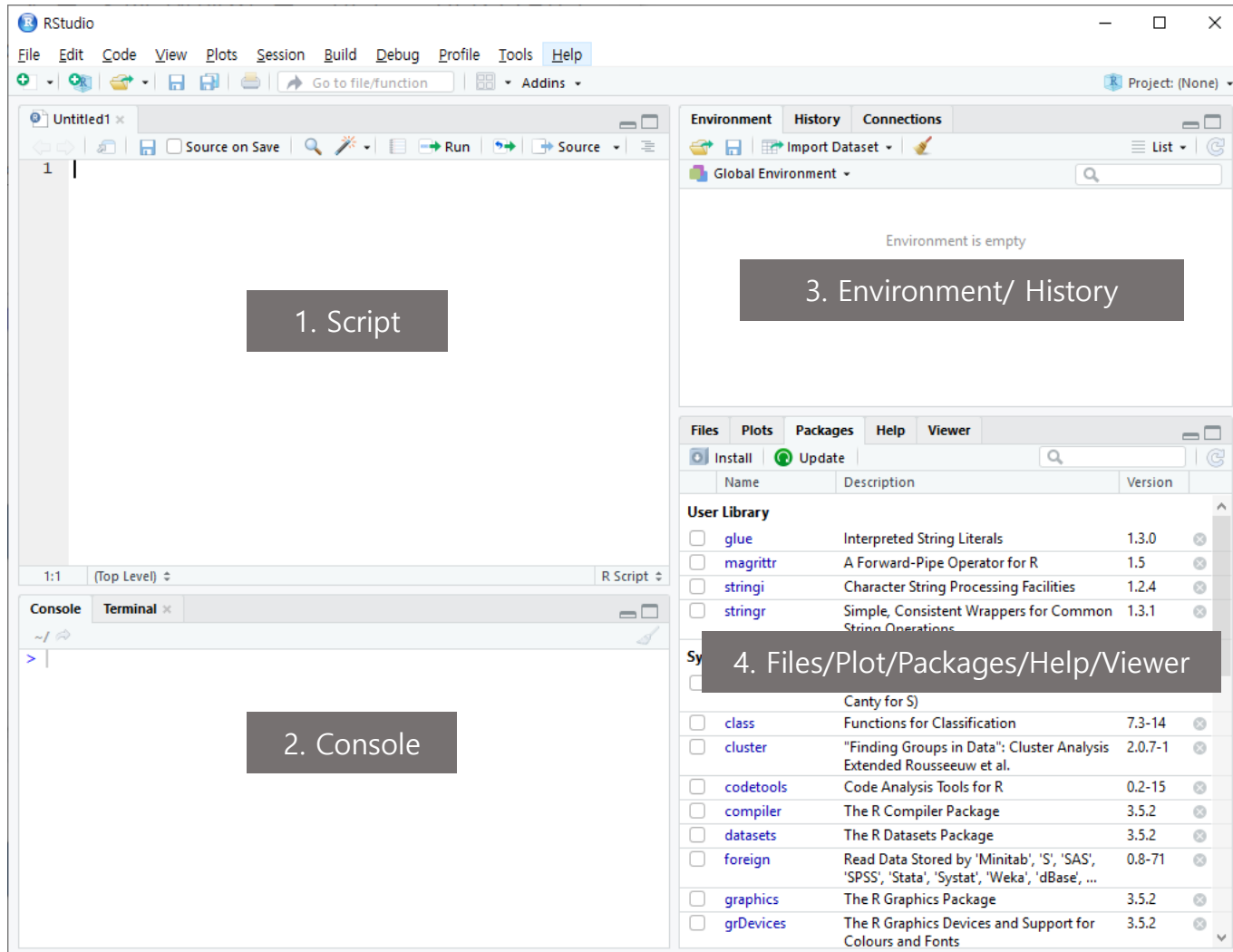


The screenshot shows the RStudio application window. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The main console area displays the R version 3.4.3 (2017-11-30) and copyright information. The right-hand pane is divided into three sections: Environment (showing 'Global Environment'), History, and Connections. Below these is a 'User Library' section with a list of installed R packages, including colorspace, dichromat, digest, ggplot2, gtable, labeling, lazyeval, magrittr, munsell, plyr, R6, RColorBrewer, Rcpp, reshape2, rlang, scales, and stringi.

⑤ 계속 다음 버튼 누른 후 설치 완료

⑥ R Studio 실행해보기

# R 소개 및 설치



## 1. Script 창

- Batch 모드로 R Script를 작성하고 실행 가능
- 'Ctrl + R' or 'Ctrl + Enter' 로 실행

## 2. Console 창

- R Script 창 or Console 창에서 작성한 프로그램의 결과 볼 수 있음
- 패키지 설치, 오류 메시지 볼 수 있음

## 3. Environment/ History 창

- 데이터 셋의 이름, 관측치 개수, 변수 개수 볼 수 있음

## 4. Files/Plot/Packages/Help/Viewer 창

- 패키지, 디렉토리 및 파일 관리
- 이미지 뷰어

## 패키지 설치하고 로드하기

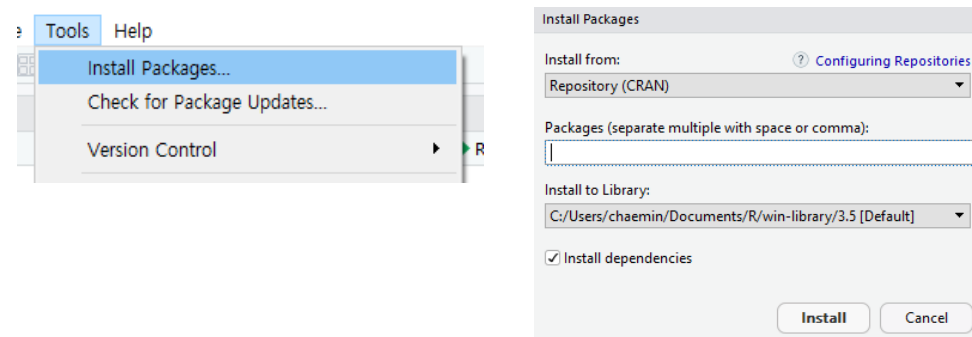
- R은 오픈소스로 수천개가 넘는 패키지를 가지고 있음
- `install.packages("패키지명")`을 쳐서 설치한 후 `library(패키지명)`으로 실행시켜주면 패키지 내의 함수를 사용할 수 있음
- `help(packages=)`를 통해 패키지를 검색할 수 있음

```
> install.packages("stringr")
Installing package into 'C:/Users/chaemin/Documents/R/win-library/3.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.5/stringr_1.4.0.zip'
Content type 'application/zip' length 214115 bytes (209 KB)
downloaded 209 KB
```

```
package 'stringr' successfully unpacked and MD5 sums checked
```

```
The downloaded binary packages are in
  C:\Users\chaemin\AppData\Local\Temp\RtmpUTYf6N\downloaded_packages
> library(stringr)
```

- Rstudio에서는 설치 이후 한번만 설정하면 됨



Tool → Install Packages → 원하는 패키지 명 입력 Install 버튼을 누름  
패키지 설치하면 항상 Library 호출을 해줘야 함



## • 할당연산자를 이용하여 변수 생성하기

- R에서 객체(object)의 형태 `name <- object(x)` #특정 객체(object)를 name에 할당

연산자	설명	예시
<code>&lt;-</code> <code>=</code> <code>&lt;&lt;-</code>	<p>왼쪽에 있는 변수에 값을 할당</p> <p>변수 &lt;- 수치, 계산식, 함수, 데이터 등</p> <p>* &lt;-를 주로 사용</p>	<pre>&gt; x1 &lt;- 2 &gt; x2 &lt;&lt;- c(1,2,3) &gt; x3 = c("a","b","c") &gt; x1 [1] 2 &gt; x2 [1] 1 2 3 &gt; x3 [1] "a" "b" "c"</pre>
<code>-&gt;</code> <code>-&gt;&gt;</code>	<p>오른쪽에 있는 변수에 값을 할당</p> <p>수치, 계산식, 함수, 데이터 등 -&gt; 변수</p>	<pre>&gt; c("a","b","c") -&gt; y1 &gt; 2 -&gt;&gt; y2 &gt; y1 [1] "a" "b" "c" &gt; y2 [1] 2</pre>

## • 변수 지정 특징

- ① 변수를 지정할 때 대소문자를 명확히 구분해야 한다
  - ② 변수명에는 영문자, 숫자, 마침표(.), 언더라인(\_)을 사용할 수 있다
  - ③ 변수명을 숫자나 숫자로 시작하는 이름으로 지정할 수 없다
  - ④ R 시스템에서 사용하는 이름은 변수명으로 사용할 수 없다 (예 : if, else, NULL, NA 등)
- 
- ① ↑ 방향키로 이전에 했던 작업 수행 가능
  - ② `q()` 사용하면 R 종료
  - ③ `#` : 주석
  - ④ R studio 스크립트 창에서 코드 입력 후 해당 부분을 선택하여 `Ctrl + Enter` 로 실행



객체를 변수로 저장하고, 출력하는 방법

Script창에서 작성해 Console 창에 결과를 띄어 줄 수 있음 : Ctrl+Enter

```
a<-3  
b=4  
c=7  
d=a+b
```

```
print(a)  
print(d)
```

따옴표, 쌍따옴표 : ', " 로 문자 처음과 끝

```
print("Hello")
```

Script창에서 작성해 Console 창에 결과를 띄어 줄 수 있음 : Ctrl+Enter

연산자	설명	예제	결과
+	더하기	3 + 2	5
-	빼기	3 - 2	1
*	곱하기	3 * 2	6
/	나누기	3 / 2	1.5
<b>^ or **</b>	승수	3 ^ 2	9
<b>x %% y</b>	X 를 y 로 나눈 나머지 값 반환	3 %% 2	1
<b>x %/% y</b>	나누기의 결과를 정수로	3 %/% 2	1

Script창에서 작성해 Console 창에 결과를 띄어 줄 수 있음 : Ctrl+Enter

연산자	설명	예제	결과
<	좌변이 보다 작은	5 < 5	FALSE
<=	좌변 이하	5 <= 5	TRUE
>	좌변이 보다 큰	5 > 5	FALSE
>=	좌변 이상	5 >= 5	TRUE
==	값이 같은	5 == 5	TRUE
!=	값이 다른	5 != 5	FALSE
!x	부정형 연산	!TRUE	FALSE
x   y	x OR y	TRUE   FALSE	TRUE
x & y	x AND y	TRUE & FALSE	FALSE
isTRUE(x)	X 의 TRUE 여부조사	isTRUE(TRUE)	TRUE

Script창에서 작성해 Console 창에 결과를 띄어 줄 수 있음 : Ctrl+Enter

- 기본 연산자

(1) 산술연산자 예시

```
> 1+2+3+4+5+6+7+8+9+10 #합
[1] 55
> (1+2+3+4+5+6+7+8+9+10)/10 #평균
[1] 5.5
> 1+2+3+4+5 #합
[1] 15
> (1+2+3+4+5)/5 #평균
[1] 3
> (1-3)^2+(2-3)^2+(3-3)^2+(4-3)^2+(5-3)^2 #편차제곱합
[1] 10
> 10/4 #분산
[1] 2.5
> 13 %/% 4 #몫
[1] 3
> 13%%4 #나머지
[1] 1
```

(2) 비교연산자 예시

```
> 3>4 #3이 4보다 큰지 판단
[1] FALSE
> 3>=4 #3이 4보다 크거나 같은지 판단
[1] FALSE
> 3<4 #3이 4보다 작은지 판단
[1] TRUE
> 3<=4 #3이 4보다 작거나 같은지 판단
[1] TRUE
> 3==4 #3이 4와 같은지 판단
[1] FALSE
> 3!=4 #3과 4가 다른지 판단
[1] TRUE
> !(3==4) #3과 4가 같다는 것을 부정
[1] TRUE
```

(3) 논리연산자 예시

```
> (5<=6)&(8<=7) #두개(또는 이상)의 조건을 동시에 만족할때 TRUE
[1] FALSE
> (5<=6)&(8>=7)
[1] TRUE
> (5<=6)|(8<=7) #두개(또는 이상)의 조건 중에서 하나만 만족해도 TRUE
[1] TRUE
> (5<=6)|(8>=7)
[1] TRUE
> (5==6)|(8>=7)
[1] TRUE
> (5==6)|(8==7)
[1] FALSE
```

R의 기본 자료형 : 자료 저장을 위해 R에서 사용하는 기본 자료형은 하나의 객체이며 그 종류는 다음과 같음

- 기본 데이터 형태

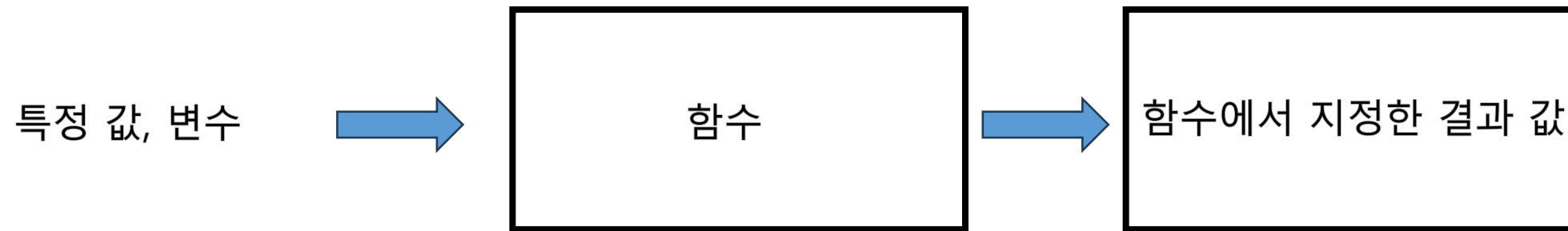
① 숫자형 (numeric) : 숫자로만 이루어진 것으로 벡터 안에 문자가 들어있으면 전체가 문자형으로 바뀜

② 문자형(character) : 문자를 처리할 때는 (" ") , ( ' ' ) 로 표시함

③ 논리값 (logic value : TRUE, FALSE) : 내부적으로 TRUE는 1, FALSE는 0 의미함

④ NA형 & NULL형 : NA형은 연산 대상에 포함, NULL형은 연산 대상에서 제외됨

함수 : 변수는 자료를 저장하는 역할을 한다면 함수는 행위를 나타냄



함수 : 변수는 자료를 저장하는 역할을 한다면 함수는 행위를 나타냄

```
함수명 <- function(인자1, 인자2, ...){  
    실행될 프로그램  
    return(반환 값) #생략가능  
}
```

반복작업이 있을 때 프로그래밍을 간결하고 깔끔하게 만들어줌

함수	내용
sin()	$\sin x$
cos()	$\cos x$
tan()	$\tan x$
sinh()	$\sinh x$
exp()	$e^x$
log()	$\log$
log(x, base=a)	$\log_a^x$
logp()	$\log_2^{(1+x)}$
sqrt()	$\sqrt{x}$
abs()	$ x $
round(x, digits)	반올림
floor()	내림
ceiling()	올림
min()	최소
max()	최대
sum()	합
choose()	C
factorial()	n!
which.max()	벡터내의 최대값 인덱스
which.min()	벡터내의 최소값 인덱스