

[R 코딩 & 데이터 분석]

Chapter 01. R 비긴즈

목차

1. 프로그래밍이란 무엇인가요?
2. R 언어에 대해 알아봅니다
3. R과 R 스튜디오를 설치해봅니다

01

프로그래밍이란 무엇인가요?

01. 프로그래밍이란 무엇인가요?

I. 프로그래밍을 배우는 이유

- **컴퓨팅 사고(Computational Thinking)** : 사람이 컴퓨터가 문제를 분석하고 해결하는 방식처럼 생각하는 것으로, 복잡한 문제를 단순화하고 이를 논리적이면서 효율적으로 해결하는 사고 과정
- **프로그래밍(programming)** : 컴퓨터가 처리할 수 있도록 문제 해결 절차를 체계적으로 서술하는 과정이며, 컴퓨터에서 실행할 소프트웨어를 만드는 과정

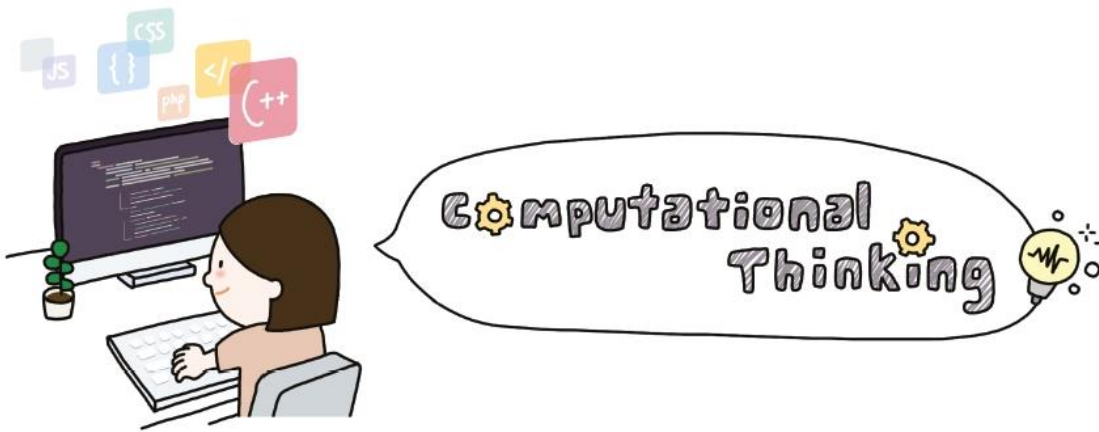


그림 1-1 컴퓨팅 사고를 통한 프로그래밍

01. 프로그래밍이란 무엇인가요?

II. 컴퓨터의 구성 : 하드웨어와 소프트웨어

- **하드웨어(hardware)** : 만질 수 있고 눈에 보이면서 형태를 가진 기계 장치들로, 컴퓨터 본체 외부에 있는 키보드, 모니터, 마우스 등과 컴퓨터 본체 내부에 있는 중앙처리장치(CPU), 메모리, 하드디스크 등이 대표적인 하드웨어
- **소프트웨어(software)** : 하드웨어의 기능을 원활하게 수행하기 위한 '명령들의 집합'으로, 엑셀, 파워포인트, 카카오톡, 페이스북 등이 이에 속함

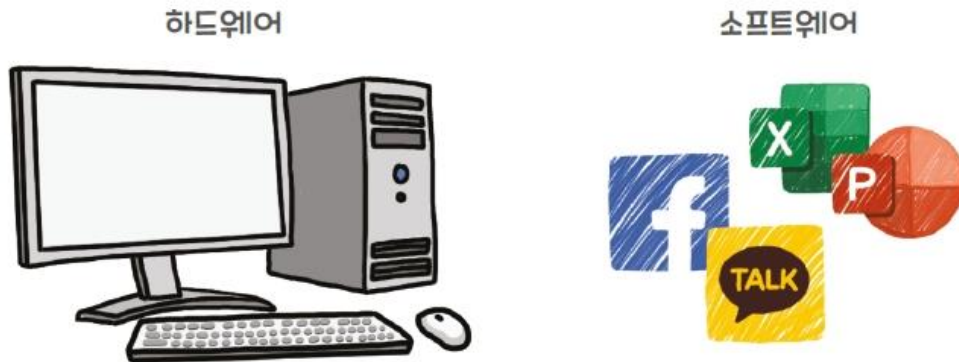


그림 1-2 하드웨어와 소프트웨어

01. 프로그래밍이란 무엇인가요?

II. 컴퓨터의 구성 : 하드웨어와 소프트웨어

- 동일한 하드웨어를 통해 다양한 일을 할 수 있는 것은 소프트웨어 덕분
- 컴퓨터에 어떤 소프트웨어가 설치되어 있는가에 따라 컴퓨터의 용도가 달라짐
- 소프트웨어는 작업의 실행 과정에서 CPU나 메모리와 같은 하드웨어 자원을 필요로 함
- 프로그래밍을 할 수 있는 능력을 갖추면 원하는 소프트웨어를 만들 수 있음

01. 프로그래밍이란 무엇인가요?

III. 프로그래밍과 프로그래밍 언어

- 프로그래밍은 요리하는 과정에 비유할 수 있으며, [그림 1-3]과 같이 정해진 절차에 따라 수행해야 함

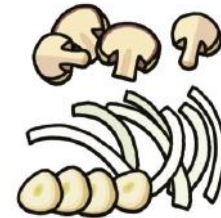
① 끓는 물에 소금 IT를 넣고 스파게티면을 10분 정도 끓인다.



② 새우는 내장을 제거하고 오징어는 흐르는 물에 깨끗이 씻어 원형으로 슬라이스한다.



③ 마늘은 편 썰고, 양파, 양송이는 슬라이스한다.



④ 달군 팬에 ③과 방울토마토를 볶은 뒤, 해물과 청주를 넣고 볶는다.



⑤ ④에 토마토소스를 넣고 끓인다.



⑥ ⑤가 끓어오르면 삶은 스파게티면을 넣고 소스를 졸여낸다.



그림 1-3 해산물 토마토 파스타 만들기

01. 프로그래밍이란 무엇인가요?

III. 프로그래밍과 프로그래밍 언어

- 컴퓨터에게 어떤 작업을 시키려면 요리의 레시피처럼 컴퓨터가 해야 할 작업의 순서와 방법을 논리적으로 제시해야 하는데, 이 과정이 프로그래밍임.
- 일상적인 언어로 컴퓨터에게 작업을 지시할 수 있으면 좋겠지만 인간의 언어는 모호성이 많이 내포되어 있음.
- 컴퓨터가 이해하기 어려워 실행할 수 없는 경우가 대부분임.

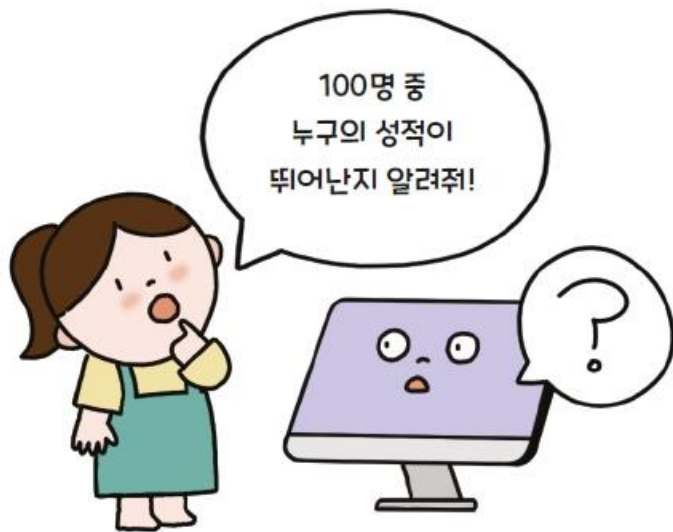


그림 1-4 일상의 언어로 명령을 내리면 컴퓨터는 이해할 수 없습니다.

01. 프로그래밍이란 무엇인가요?

III. 프로그래밍과 프로그래밍 언어

- 프로그래밍 언어(programming language) :
컴퓨터가 어떤 작업을 수행하기 위한 프로그램을 작성하는 데 사용하는 언어
- [그림 1-5]와 같이 간단한 영어 단어와 기호들로 구성되어 있으며 문법도 존재함
- 프로그래밍을 공부한다는 것은 프로그래밍 언어의 문법 구조를 배우고 이것을 활용하여 컴퓨터에게 일을 시키기 위한 '작업 지시서'를 작성하는 방법, 즉 프로그램 작성 방법을 학습하는 것

```
sum <- 0
for (i in 1:1000) {
  sum <- sum + i
}
print(sum)
```

그림 1-5 숫자 1~1,000까지의 합계를 구하는 간단한 프로그램

01. 프로그래밍이란 무엇인가요?

III. 프로그래밍과 프로그래밍 언어

- 대표적인 프로그래밍 언어 : C, 자바, 파이썬 등
- 소프트웨어를 만들고자 한다면 이들 언어 중 하나를 반드시 배워야만 함

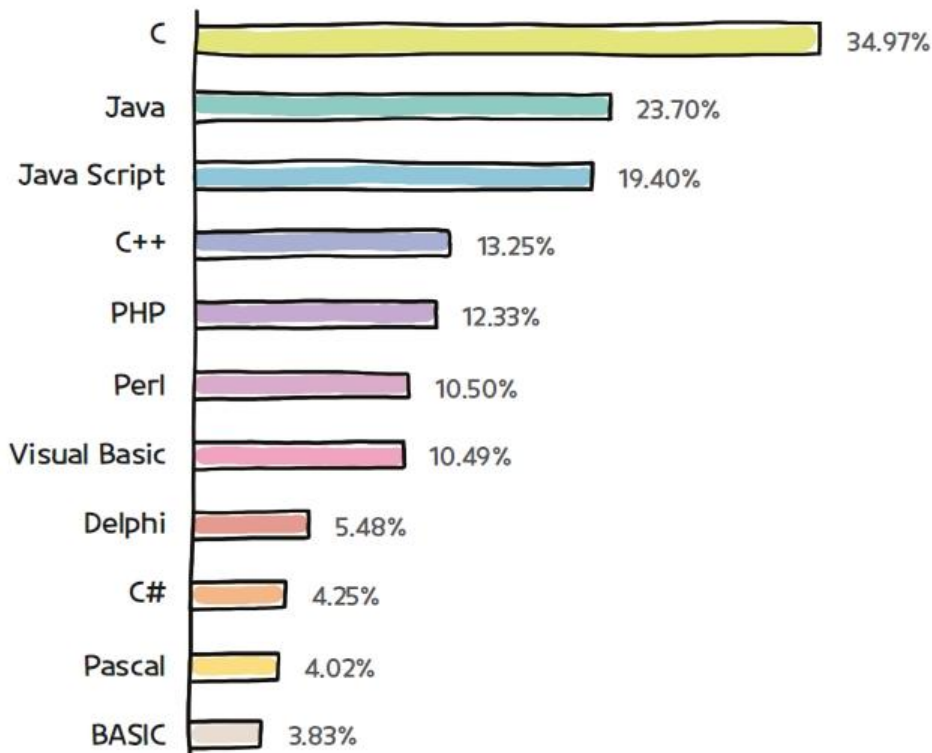


그림 1-6 프로그래밍 언어 ©Popular programming languages 1965-2019

01. 프로그래밍이란 무엇인가요?

III. 프로그래밍과 프로그래밍 언어

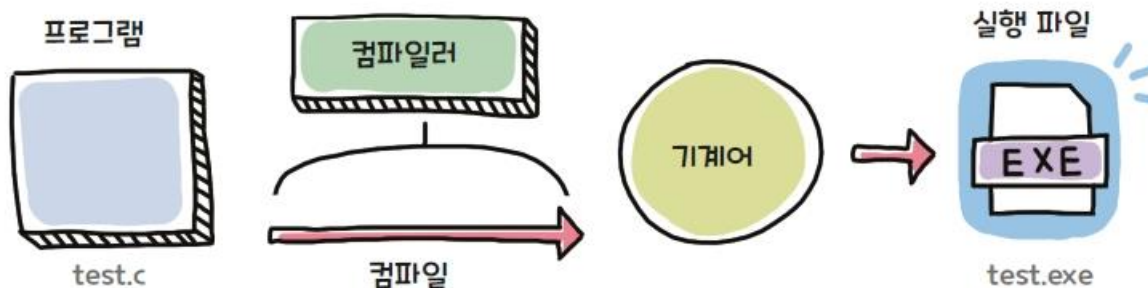
- **프로그램(program)** : 특정 프로그래밍 언어를 사용하여 문제 해결 과정을 작성한 문서로, 컴퓨터에게 일을 시키기 위한 '작업 지시서'
- **기계어(machine language)** : 컴퓨터는 이해할 수 있는 언어로, 0과 1로 구성됨
- **컴파일(compile)** : 인간의 언어에 가까운 프로그램을 컴퓨터가 이해할 수 있는 기계어 형태로 변환하는 과정
- **컴파일러(compiler)** : 컴파일을 하기 위해 사용하는 특별한 소프트웨어
- **실행 파일(executable file)** : 프로그램을 컴파일하면 생성되는 컴퓨터가 이해하여 실행 가능한 기계어 파일

01. 프로그래밍이란 무엇인가요?

III. 프로그래밍과 프로그래밍 언어

1000110	0101010	0011100	1101101
1000101	0101001	0011111	1101110
1000011	0101111	0011001	1101000
1001111	0100011	0010101	1100100
1010111	0111011	0001101	1111100

(a) 기계어의 예



(b) 컴파일을 통한 실행 파일이 생성

그림 1-7 프로그램이 만들어지는 과정

01. 프로그래밍이란 무엇인가요?

III. 프로그래밍과 프로그래밍 언어

하나 더 알기

실행 파일

- 우리가 컴퓨터의 탐색기를 통해 특정 폴더의 내용을 보았을 때 파일의 확장자가 'exe'인 파일이 바로 실행 파일입니다. 실행 파일을 더블클릭하면 그 안에 있는 기계어로 된 명령문들이 실행되면서 소프트웨어가 작동됩니다.

01. 프로그래밍이란 무엇인가요?

III. 프로그래밍과 프로그래밍 언어

- 많은 사람들에게 이러한 프로그래밍 과정은 매우 어려운 것으로 다가올 수 있음
- 컴퓨터 비전공자들도 쉽게 배워서 업무에 사용할 수 있는 프로그래밍 언어가 필요했고, 그러한 목적으로 개발된 언어 중의 하나가 바로 본 교재를 통해서 배우고자 하는 R 언어임

02

R 언어에 대해 알아보니다

02. R 언어에 대해 알아봅니다

I. R 언어의 특징

- R 언어는 비교적 최근에 개발된 프로그래밍 언어 중 하나
- R은 1993년, 뉴질랜드 오클랜드 대학의 로스 이하카(Ross Ihaka)와 로버트 젠틀맨(Robert Gentleman)에 의해 통계 프로그래밍 언어인 S-PLUS의 무료 버전 형태



그림 1-8 R 언어 로고

02. R 언어에 대해 알아봅니다

I. R 언어의 특징

■ 데이터 분석에 특화된 언어

- R은 통계를 포함한 데이터 분석 작업에 활용할 목적으로 개발된 언어
- R은 컴파일 과정 없이도 바로 실행하여 결과를 확인할 수 있음
- R로 작성한 것을 '프로그램'이 아니라 스크립트(script)라고 부름

■ 탄탄한 사용자 커뮤니티

- R은 사용자층이 두텁기 때문에 다양한 커뮤니티가 존재함
- 초보자를 위한 학습 자료도 풍부하게 존재함
- 국내 검색 사이트를 통해서 찾을 수 있는 한글 자료들이 증가하는 추세

02. R 언어에 대해 알아봅니다

I. R 언어의 특징

The screenshot shows the homepage of R-bloggers, a website dedicated to R news and tutorials. The header features the R-bloggers logo and the tagline "R news and tutorials contributed by hundreds of R bloggers". Navigation links include HOME, ABOUT, RSS, ADD YOUR BLOG!, LEARN R, RJOBS, and CONTACT US. A search bar is located on the right side. The main content area is titled "Tutorials for learning R" and includes a post from December 10, 2015, by Tal Galili. The post content is partially visible, showing R code and a scatter plot. The right sidebar contains a search bar, a subscribe button, and a list of most viewed posts.

R-bloggers
R news and tutorials contributed by hundreds of R bloggers

HOME ABOUT RSS ADD YOUR BLOG! LEARN R RJOBS CONTACT US

Tutorials for learning R

Posted on December 10, 2015 by Tal Galili in R guest posts | 0 Comments

Instructions

- 1. Import the dataset of the first code chunk on the right, then the dataset of the second code chunk on the left and use the instructions to create the plot.
- 2. The first chunk of code sets up the data and the second chunk creates the plot. The third chunk shows the final result of the plot.
- 3. Add a point to the dataset and see the effect on the plot.
- 4. Add a legend to the plot.
- 5. Add a title to the plot.
- 6. Add a subtitle to the plot.
- 7. Add a label to the plot.
- 8. Add a color to the plot.
- 9. Add a shape to the plot.
- 10. Add a size to the plot.

Most viewed posts (weekly)

그림 1-9 대표적 R 커뮤니티 R-bloggers(<https://www.r-bloggers.com/>)

02. R 언어에 대해 알아보니다

I. R 언어의 특징

■ 다양한 패키지 제공

- R은 데이터 분석에 사용되는 함수들을 종류별로 묶어 패키지 형태로 제공
- 데이터 분석에 필요한 거의 모든 기능이 제공
- 최신 이론이 발표되면 바로 R 패키지가 만들어지기 때문에 신속하게 최신 이론을 데이터 분석에 활용하는 것이 가능

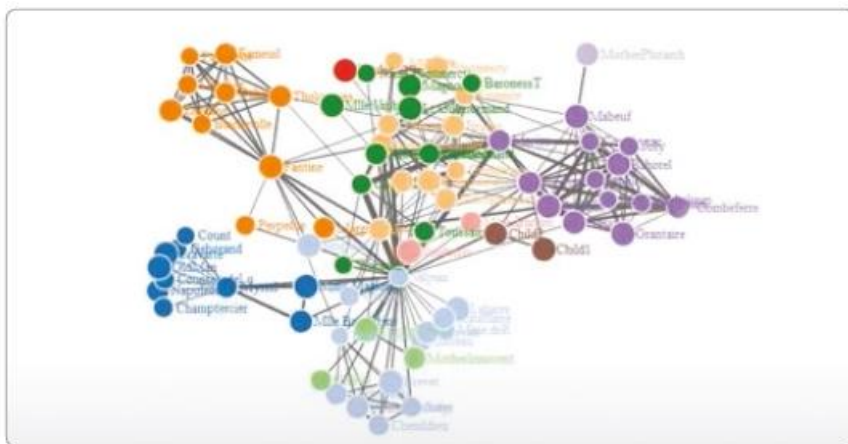


그림 1-10 networkD3 패키지 : 네트워크 형태의 데이터를 시각화하는 데 사용됩니다.

02. R 언어에 대해 알아봅니다

I. R 언어의 특징

■ 미적이고 기능적인 통계 그래프 제공

- 데이터 분석에 있어서 분석 결과를 시각적으로 표현하는 것은 매우 중요함
- R에서는 ggplot이라는 패키지를 통해 아름다우면서도 기능적인 그래프를 쉽게 작성할 수 있도록 지원함

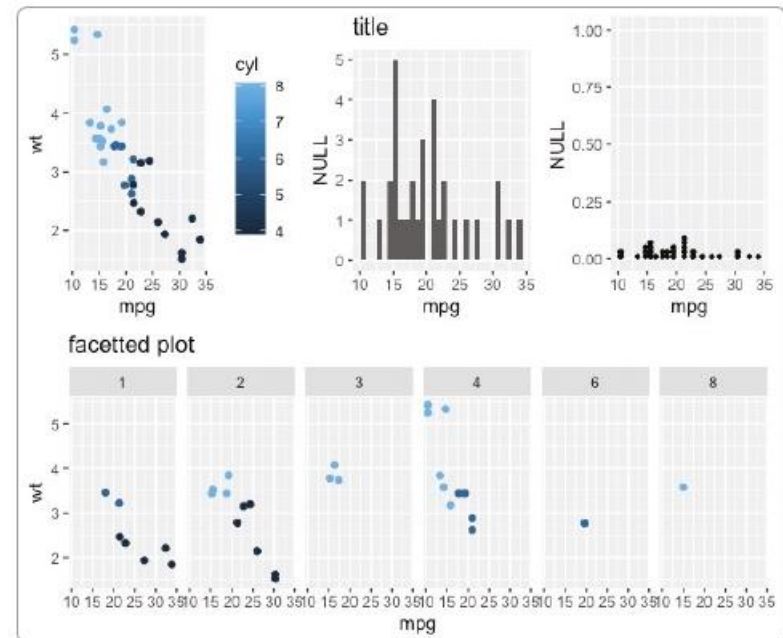


그림 1-11 ggplot 그래프의 예

02. R 언어에 대해 알아봅니다

I. R 언어의 특징

■ 편리한 프로그래밍 환경

- R 프로그래밍을 위한 통합 개발 환경으로 R 스튜디오가 제공되어 모든 작업을 R 스튜디오 내에서 처리할 수 있음
- 프로그램 작성·실행·수정 등 여러 작업을 수행하려면 보다 편리한 작업 환경이 필요함
- 이러한 작업 환경을 통합 개발 환경, IDE(Integrated Development Environment)라고 함

02. R 언어에 대해 알아봅니다

I. R 언어의 특징

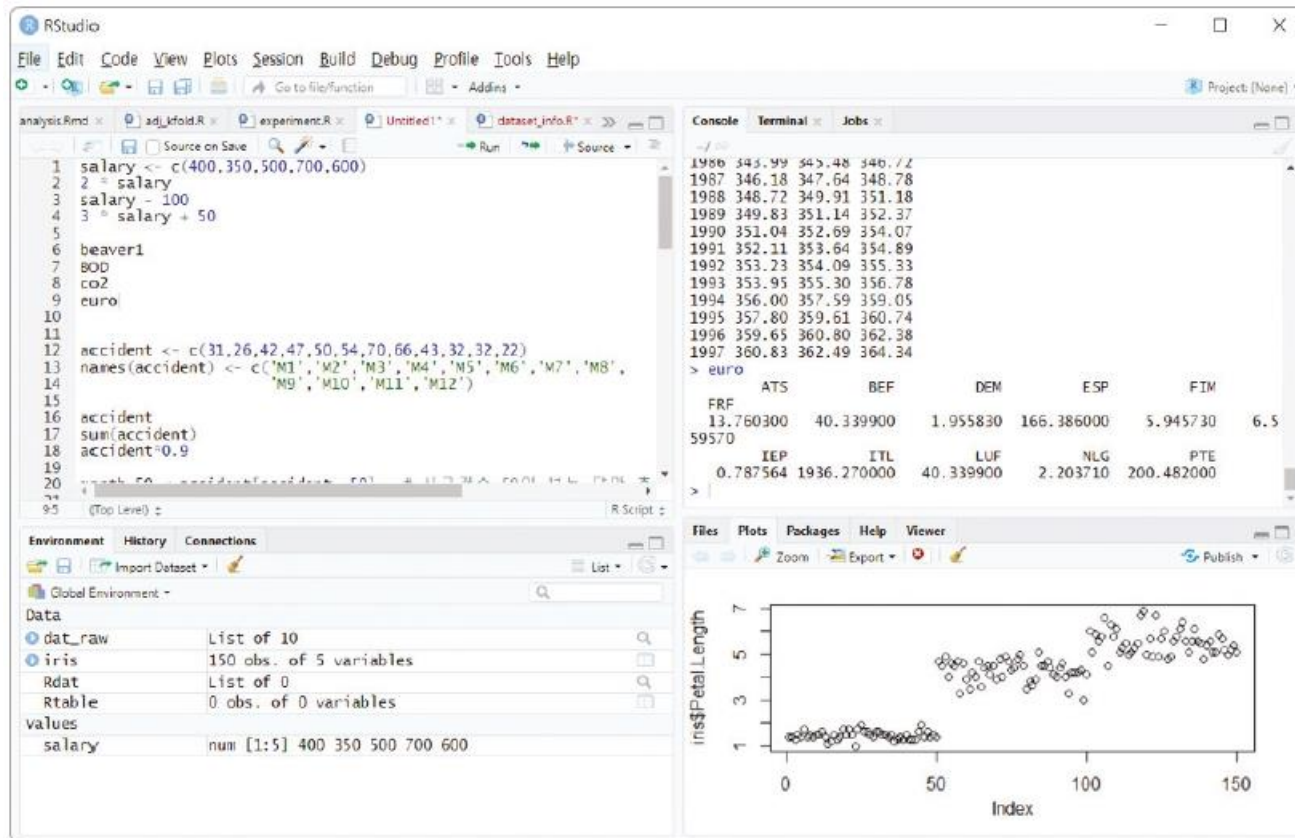


그림 1-12 R 스튜디오 실행 화면

02. R 언어에 대해 알아봅니다

I. R 언어의 특징

■ 무료 사용

- R은 무료로 사용할 수 있는 오픈소스(open source) 소프트웨어
- 1년에 1~2차례 정기적으로 업데이트가 이루어져 지속적으로 기능이 향상
- R은 윈도우 환경뿐만 아니라 리눅스, 맥OS 환경에서도 설치 및 사용이 가능

02. R 언어에 대해 알아봅니다

II. R을 배우는 이유

- **4차 산업혁명(The Fourth Industrial Revolution) :**

인공지능, 빅데이터, 로봇, 사물인터넷 등 새로운 과학 기술이 정치·경제·문화 전반에 영향을 미치면서 이러한 변화를 수용하고 발전 가능성을 최대화시키는 시대로, 인공지능(AI, Artificial Intelligence)과 빅데이터(Big Data)가 핵심 기술로 인식됨

- 4차 산업혁명의 중심은 '데이터'
- 시대적 흐름은 데이터를 잘 다룰 줄 아는 기업과 개인이 경쟁에서 우위를 점할 수 있음을 의미임
- 컴퓨팅 사고와 데이터 활용 능력은 전공을 불문하고 어떤 분야로 진출하든 점점 더 중요한 스펙임

02. R 언어에 대해 알아봅니다

II. R을 배우는 이유

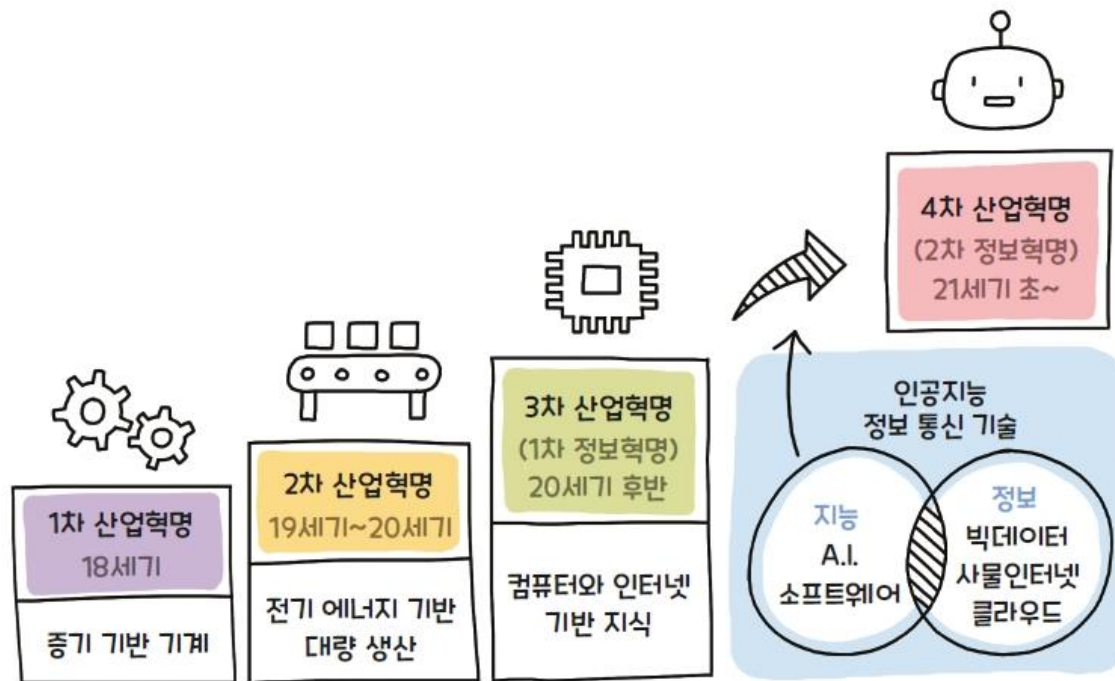


그림 1-13 산업혁명의 발전과 핵심 기술

- R은 배우기 쉬우면서도 강력한 데이터 처리 및 데이터 분석 능력을 제공함
- 프로그래밍 언어로 컴퓨팅 사고를 배우기에도 적합함

03

R과 R 스튜디오를 설치해봅시다

02. R과 R 스튜디오를 설치해봅니다

- R 프로그래밍을 학습하기 위해서는 소프트웨어인 R과 R 스튜디오(RStudio)를 설치
- R은 프로그램을 작성하고 실행하기 위한 소프트웨어이고, R 스튜디오는 R 프로그래밍을 편리하게 작업할 수 있도록 도움을 주는 보조 소프트웨어
- R을 먼저 설치하고, 그 다음에 R 스튜디오를 설치

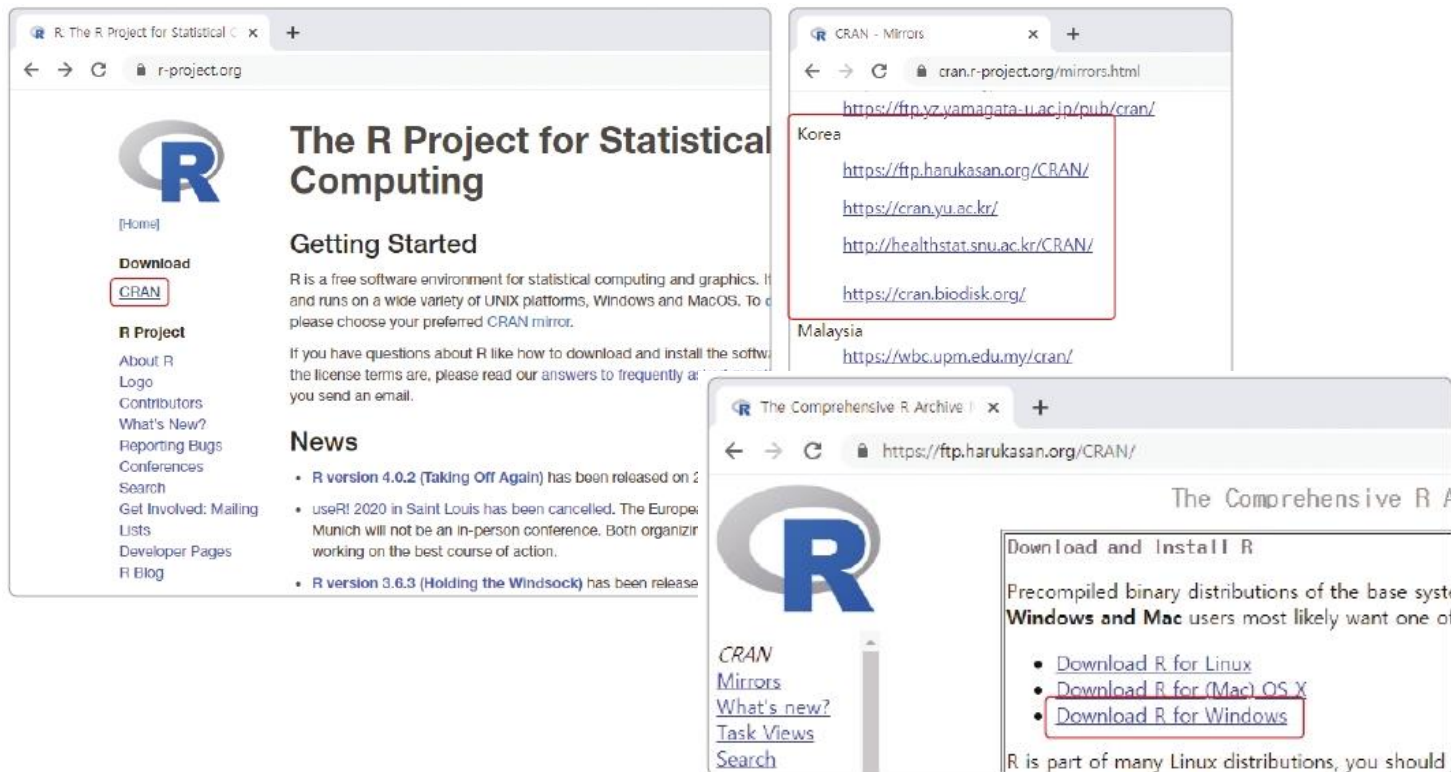
I. R 설치하기

- R을 사용하려면 먼저 [공식 웹사이트](#)에서 R을 다운로드하고 설치해야 함

02. R과 R 스튜디오를 설치해봅시다

I. R 설치하기

- 1) [Download] 항목의 [CRAN]을 클릭 → [Korea] 항목의 링크 → [Download R for Windows]를 클릭



02. R과 R 스튜디오를 설치해봅니다

I. R 설치하기

하나 더 알기

미러 사이트

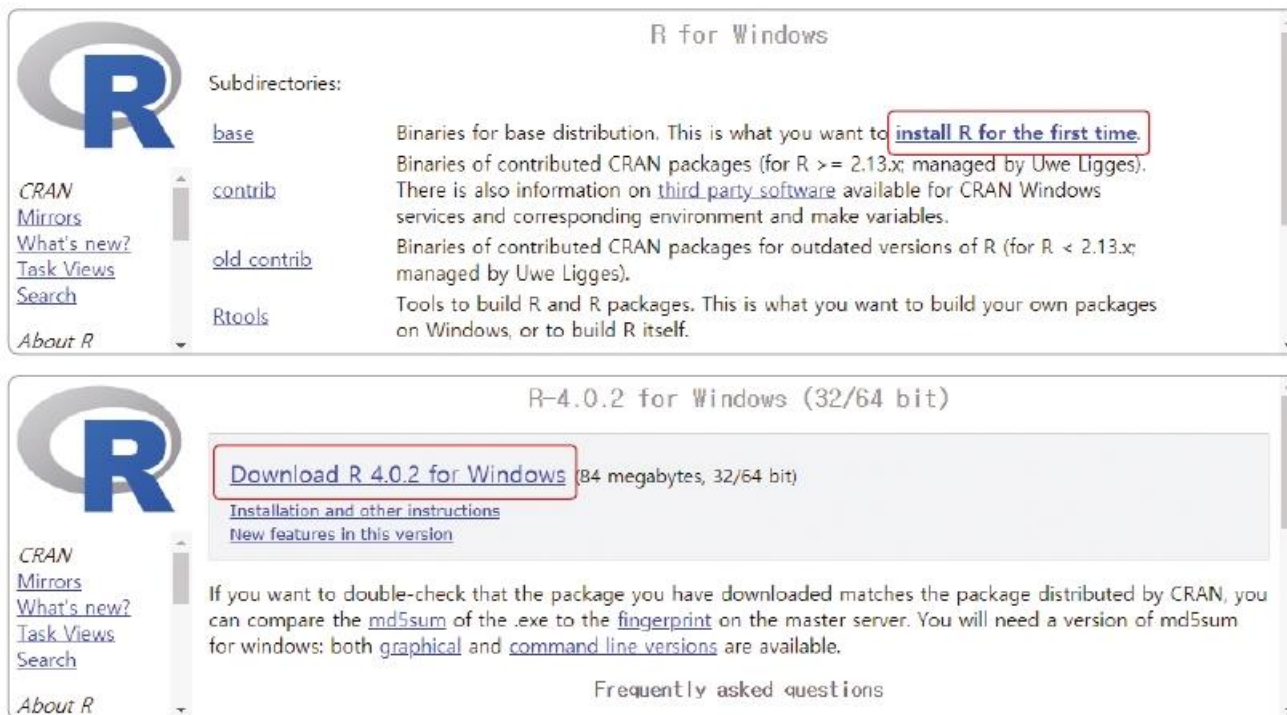
- 하나의 웹사이트에서만 R을 제공하면 전 세계 사용자가 몰려 이용 속도가 저하될 수 있습니다. 이를 방지하기 위해 동일한 사이트를 여러 곳에 복사해서 사용자들을 분산시켜야 합니다. 이렇게 원래 사이트에서 복사된 사이트를 미러 (Mirrors) 사이트라고 합니다, 사용자가 지리적으로 가까운 미러 사이트를 이용하면 보다 빠르게 파일을 다운로드받을 수 있습니다.

02. R과 R 스튜디오를 설치해봅시다

I. R 설치하기

2) [install R for the first time] 링크를 클릭

→ [Download R 4.0.2 for Windows]를 클릭



The screenshot shows two browser windows. The top window is titled 'R for Windows' and displays a list of subdirectories: base, contrib, old contrib, and Rtools. The 'base' subdirectory is selected, and the text 'Binaries for base distribution. This is what you want to [install R for the first time.](#)' is visible. The bottom window is titled 'R-4.0.2 for Windows (32/64 bit)' and displays a large button labeled 'Download R 4.0.2 for Windows (84 megabytes, 32/64 bit)'. Below this button, there are links for 'Installation and other instructions' and 'New features in this version'. The text 'If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the [md5sum](#) of the .exe to the [fingerprint](#) on the master server. You will need a version of md5sum for windows: both [graphical](#) and [command line versions](#) are available.' is also visible.

R for Windows

Subdirectories:

- [base](#): Binaries for base distribution. This is what you want to [install R for the first time.](#)
- [contrib](#): Binaries of contributed CRAN packages (for R >= 2.13.x; managed by Uwe Ligges). There is also information on [third party software](#) available for CRAN Windows services and corresponding environment and make variables.
- [old contrib](#): Binaries of contributed CRAN packages for outdated versions of R (for R < 2.13.x; managed by Uwe Ligges).
- [Rtools](#): Tools to build R and R packages. This is what you want to build your own packages on Windows, or to build R itself.

R-4.0.2 for Windows (32/64 bit)

[Download R 4.0.2 for Windows](#) (84 megabytes, 32/64 bit)

[Installation and other instructions](#)

[New features in this version](#)

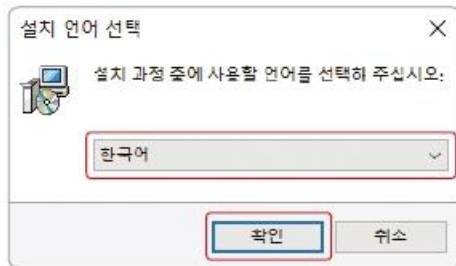
If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the [md5sum](#) of the .exe to the [fingerprint](#) on the master server. You will need a version of md5sum for windows: both [graphical](#) and [command line versions](#) are available.

Frequently asked questions

02. R과 R 스튜디오를 설치해봅니다

I. R 설치하기

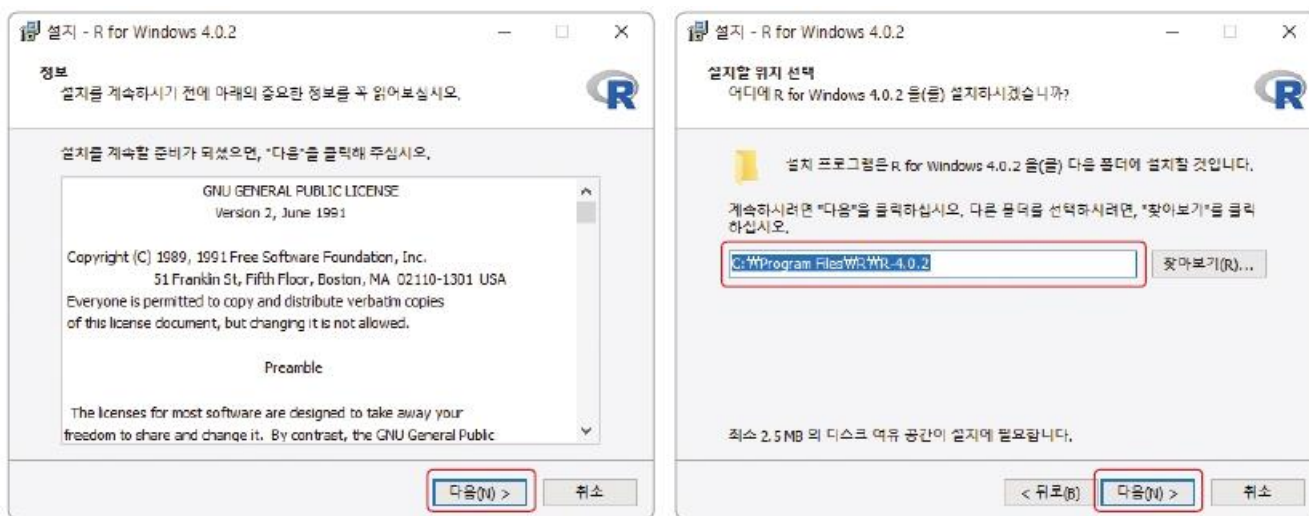
3) [한국어]를 선택하고 [확인] 버튼을 클릭



02. R과 R 스튜디오를 설치해봅니다

I. R 설치하기

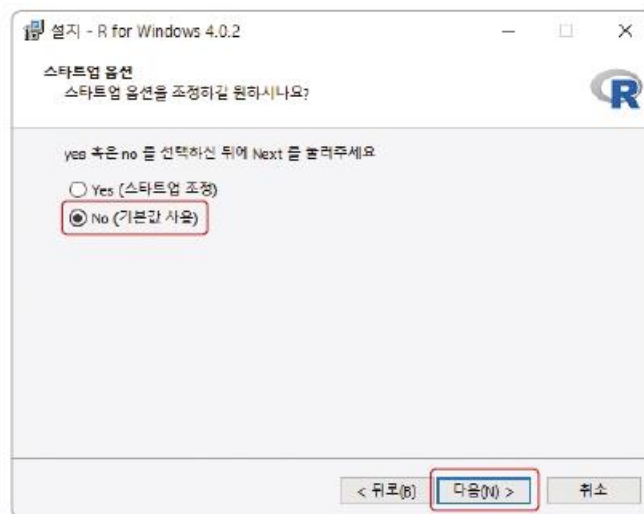
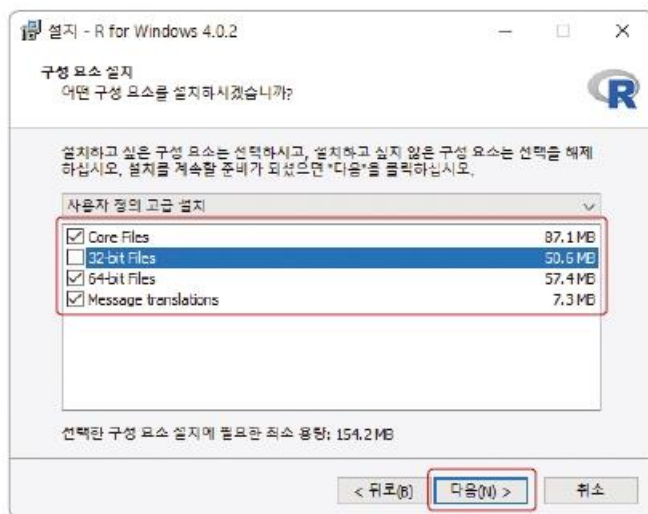
4) 설치할 위치 선택에서 경로를 선택한 후 [다음] 버튼을 클릭



02. R과 R 스튜디오를 설치해봅니다

I. R 설치하기

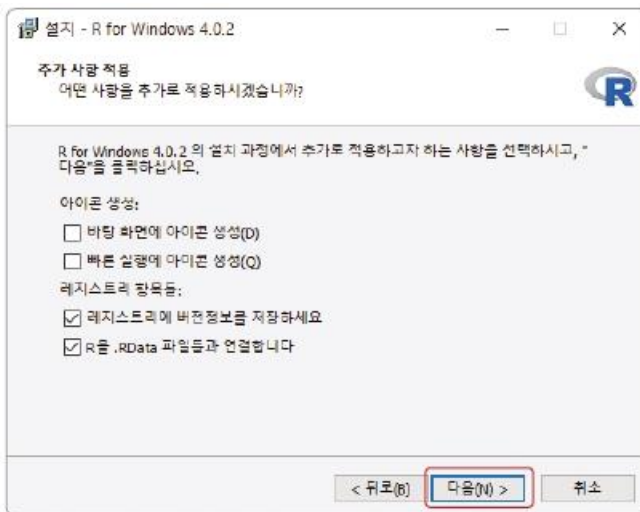
- 5) 구성 요소 설치에서 필요한 항목을 체크하고 [다음] 버튼을 클릭
→ 스타트업 옵션에서 [No]를 선택하고 [다음] 버튼을 클릭



02. R과 R 스튜디오를 설치해봅니다

I. R 설치하기

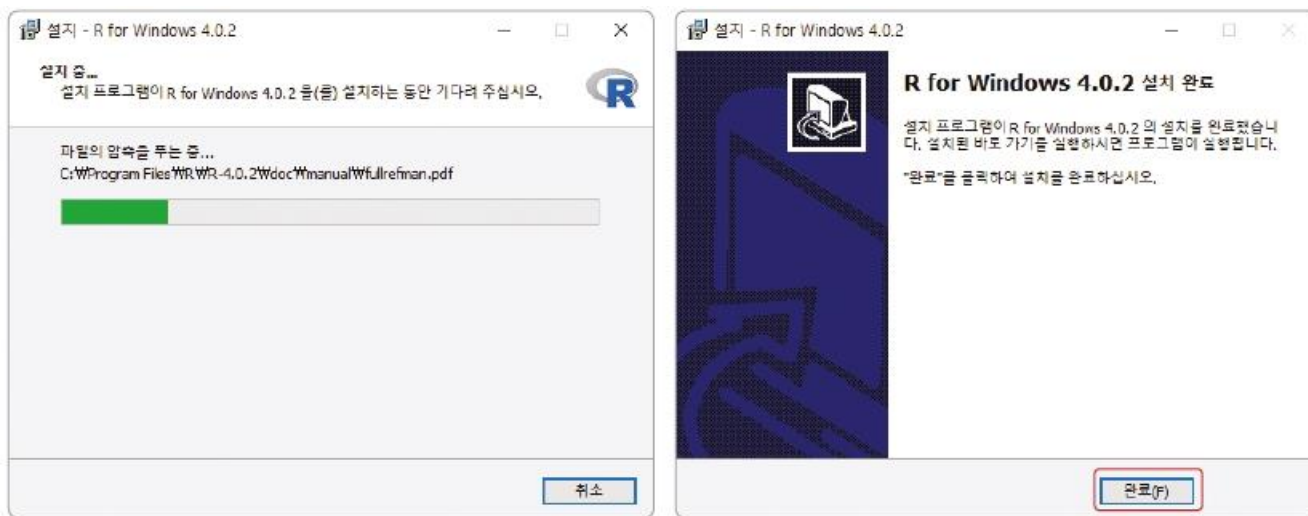
- 6) 시작 메뉴 폴더 선택은 내용 변경 없이 [다음] 버튼을 클릭
→ 추가 사항 적용도 내용 변경 없이 [다음] 버튼을 클릭



02. R과 R 스튜디오를 설치해봅니다

I. R 설치하기

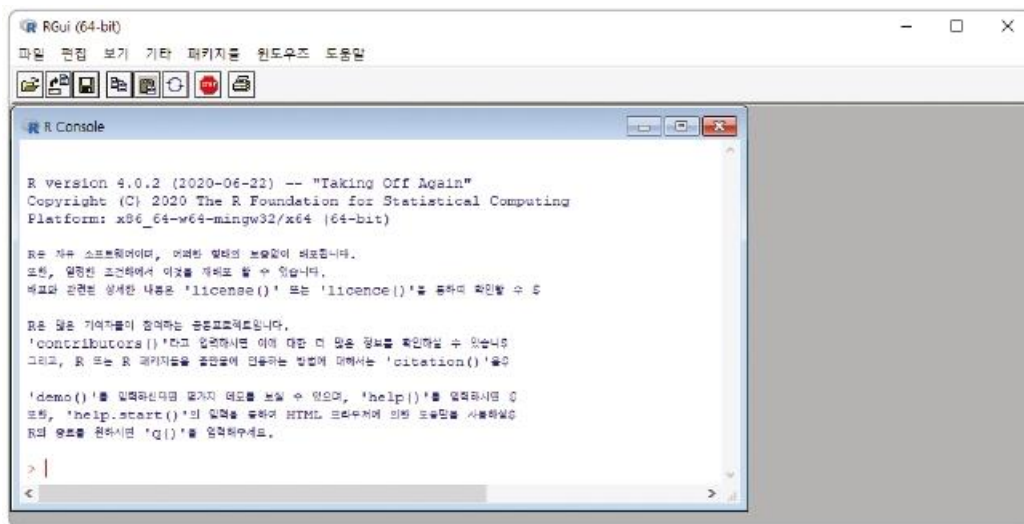
7) 설치 완료 창이 열리면 [완료] 버튼을 클릭



02. R과 R 스튜디오를 설치해봅니다

I. R 설치하기

8) 설치가 완료되면 윈도우 시작 메뉴에서 R을 클릭하여 정상적으로 실행되는지 확인함



02. R과 R 스튜디오를 설치해봅니다

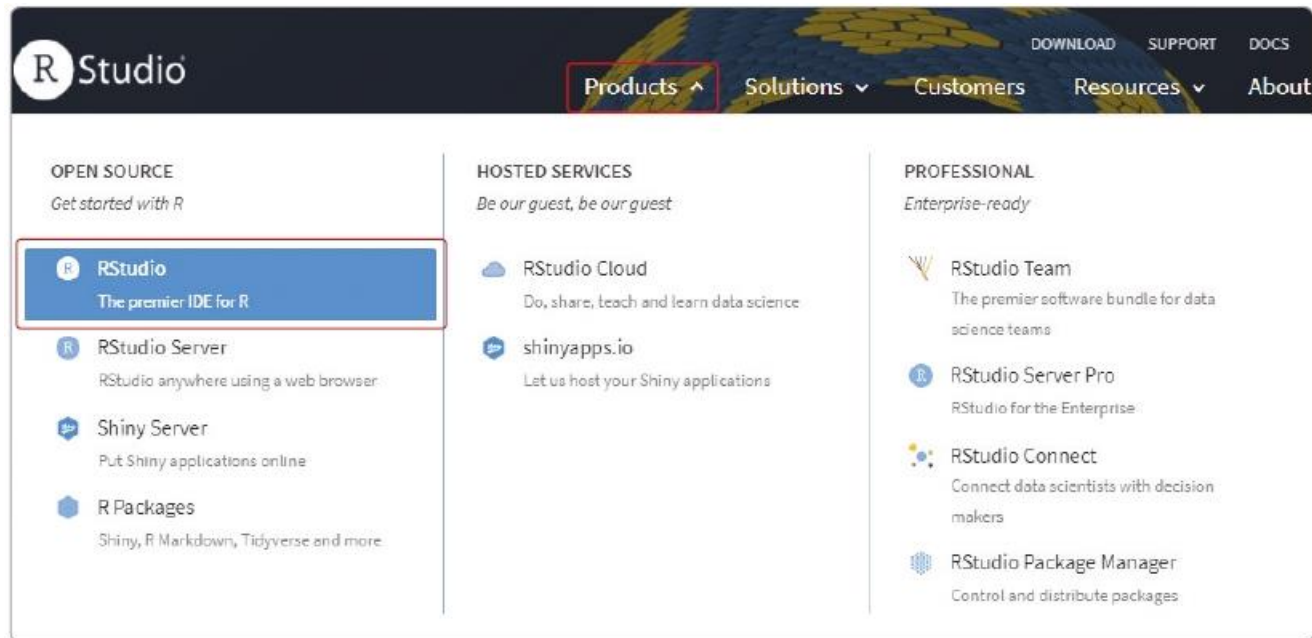
II. R 스튜디오 설치하기

- R 스튜디오(RStudio) : R을 편리하게 사용할 수 있도록 돕는 통합 개발 환경 소프트웨어
- R 스튜디오 공식 웹사이트에서 R 스튜디오를 다운로드하고 설치해야 함

02. R과 R 스튜디오를 설치해봅시다

II. R 스튜디오 설치하기


- 1) 웹사이트인 <http://www.rstudio.com/>에 접속 → [Products]를 클릭
→ OPEN SOURCE 카테고리에서 [RStudio]를 클릭



02. R과 R 스튜디오를 설치해봅니다

II. R 스튜디오 설치하기

2) [DOWNLOAD RSTUDIO DESKTOP]을 클릭

 **Studio Desktop**

	Open Source Edition	RStudio Desktop Pro
Overview	<ul style="list-style-type: none">• Access RStudio locally• Syntax highlighting, code completion, and smart indentation• Execute R code directly from the source editor• Quickly jump to function definitions• Easily manage multiple working directories using projects• Integrated R help and documentation• Interactive debugger to diagnose and fix errors quickly• Extensive package development tools	<p>All of the features of open source; plus:</p> <ul style="list-style-type: none">• A commercial license for organizations not able to use AGPL software• Access to priority support• RStudio Professional Drivers• Connect directly to your RStudio Server Pro instance remotely
Support	Community forums only	<ul style="list-style-type: none">• Priority Email Support• 8 hour response during business hours (ET)
License	AGPL v3	RStudio License Agreement
Pricing	Free	\$995/year
	DOWNLOAD RSTUDIO DESKTOP	DOWNLOAD FREE RSTUDIO DESKTOP PRO TRIAL

02. R과 R 스튜디오를 설치해봅니다


II. R 스튜디오 설치하기

3) [DOWNLOAD]를 클릭

Choose Your Version

RStudio is a set of integrated tools designed to help you be more productive with R. It includes a console, syntax-highlighting editor that supports direct code execution, and a variety of robust tools for plotting, viewing history, debugging and managing your workspace.

[LEARN MORE ABOUT RSTUDIO FEATURES](#)



RStudio's new solution for every professional data science team. RStudio Team includes RStudio Server Pro, RStudio Connect and RStudio Package Manager.

[LEARN MORE](#)

RStudio Desktop Open Source License	RStudio Desktop Commercial License	RStudio Server Open Source License	RStudio Server Pro Commercial License
Free	\$995 /year	Free	\$4,975 /year (5 Named Users)
DOWNLOAD Learn more	BUY Learn more	DOWNLOAD Learn more	BUY Evaluation Learn more
Integrated Tools for R ✓	✓	✓	✓


02. R과 R 스튜디오를 설치해봅니다

II. R 스튜디오 설치하기


4) [DOWNLOAD RSTUDIO FOR WINDOWS]를 클릭

RStudio Desktop 1.3.1073 - Release Notes

1. Install R. RStudio requires R 3.0.1+.
2. Download RStudio Desktop. Recommended for your system:

 **DOWNLOAD RSTUDIO FOR WINDOWS**
1.3.1073 | 171.62MB

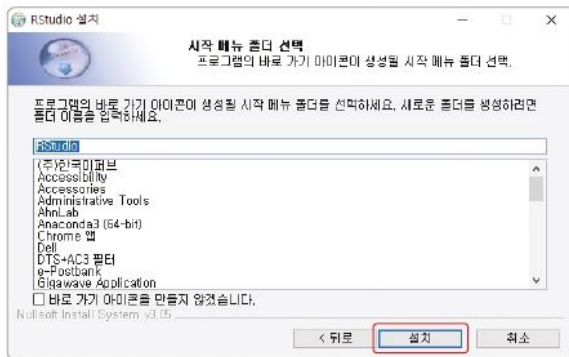
Requires Windows 10/8/7 (64-bit)



02. R과 R 스튜디오를 설치해봅시다

II. R 스튜디오 설치하기

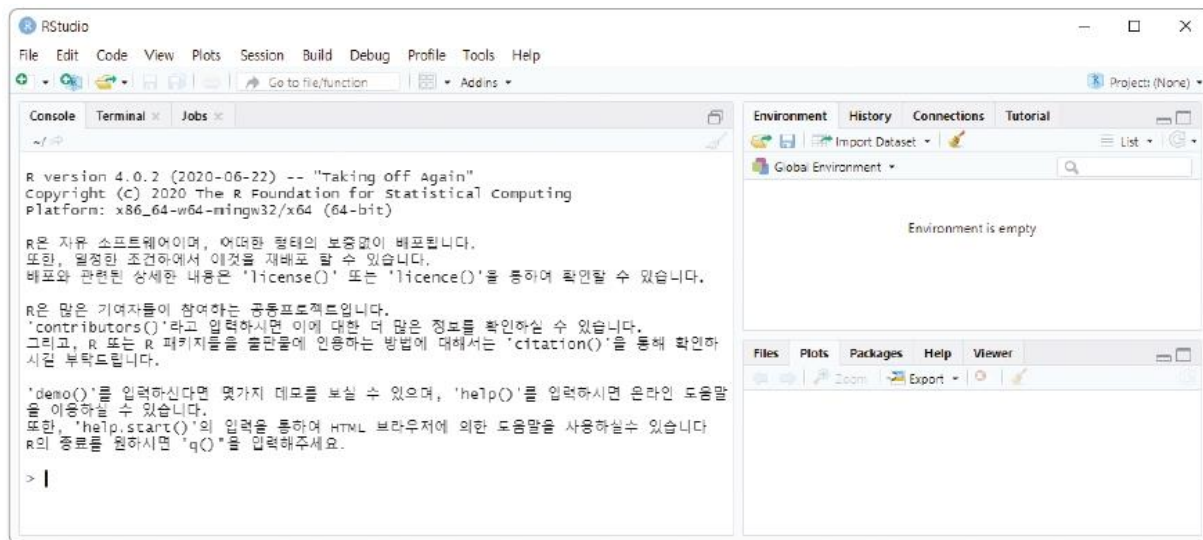
5) 설치 파일을 더블클릭 → 계속해서 [다음] 버튼을 클릭
→ [마침] 버튼을 클릭



02. R과 R 스튜디오를 설치해봅시다

II. R 스튜디오 설치하기

6) 윈도우 시작 메뉴에서 R 스튜디오를 클릭



Thank You !