

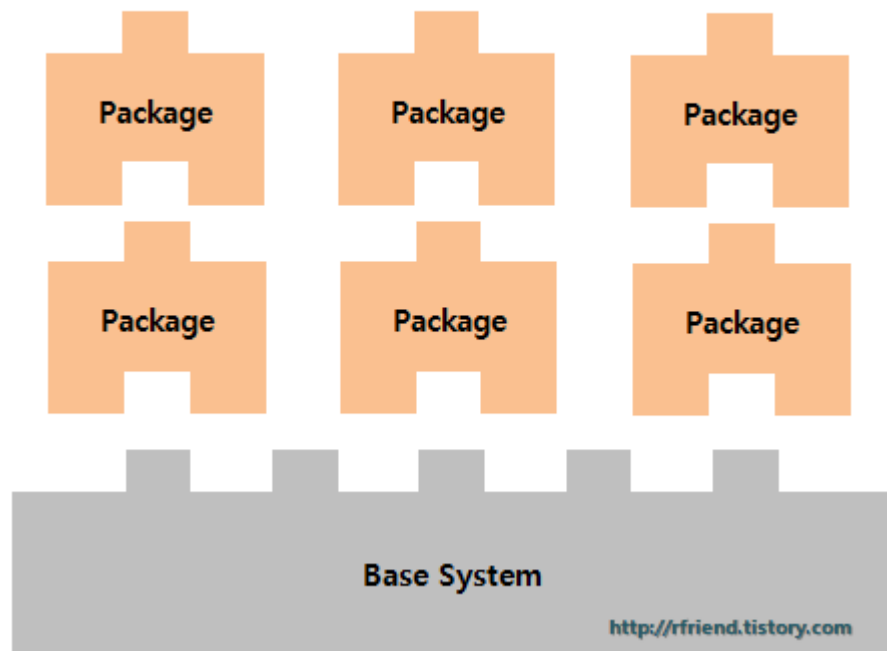
교육 빅데이터의 이해와 활용(10주차)

흥도초등학교

교사 박 정

R 패키지 이해

【 Base System과 Package의 관계 】



```
R Console (64-bit)
File Edit Misc Packages Windows Help

R version 2.15.2 (2012-10-26) -- "Trick or Treat"
Copyright (C) 2012 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: x86_64-w64-mingw32/x64 (64-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.

Natural language support but running in an English locale

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.

[Previous saved workspace restored]
> install.packages("gplots")
```

```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows [Version 10.0.17134.191]
(c) 2018 Microsoft Corporation. All rights reserved.

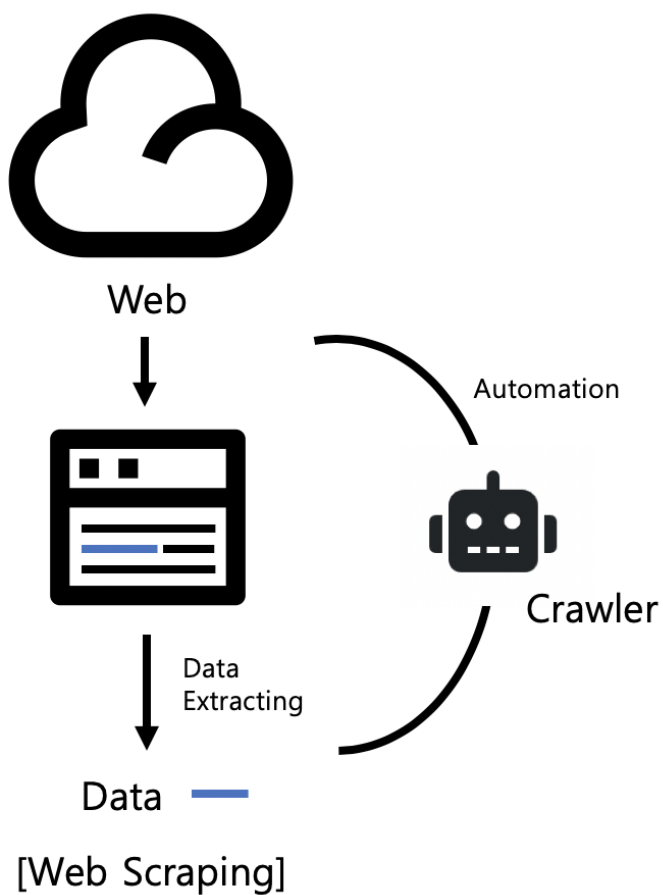
C:\Users\wtongchun>python --version
Python 3.7.0

C:\Users\wtongchun>pip --version
pip 18.0 from c:\python37\lib\site-packages\pip (python 3.7)

C:\Users\wtongchun>pip install opencv-python
Collecting opencv-python
  Downloading https://files.pythonhosted.org/packages/20/f3/90410344a0d68b9888e64b16152807c33fa5a8eaeefdb3beb18d88eea62159/opencv_python-3.4.2.17-cp37-cp37m-win_amd64.whl (33.8MB)
  100% |#####| 33.8MB 775KB/s
Collecting numpy>=1.14.5 (from opencv-python)
  Downloading https://files.pythonhosted.org/packages/8b/8a/5edea6c9759b9c569542ad4da07bba0c03ffe7cfb15d8bbe59b417e99a84/numpy-1.15.0-cp37-none-win_amd64.whl (13.5MB)
  100% |#####| 13.5MB 2.3MB/s
Installing collected packages: numpy, opencv-python
Successfully installed numpy-1.15.0 opencv-python-3.4.2.17

C:\Users\wtongchun>
```

웹 데이터 가져오기



웹 페이지가 우리 눈에 보이기 까지

1. 클라이언트가 서버에게 contents를 요청한다.
2. 서버는 요청 받은 contents를 클라이언트에게 건네준다.
3. 브라우저는 서버에게 받은 HTML을 해석하여 화면에 보여준다.



웹 데이터 가져오기

Web Scraping Using

Monitoring and tracking prices set by your competitors is vital part of the process of running an ecommerce business to keep your product price competitive and attractive.



O'REILLY®

초간단 나만의 웹 크롤러로 원하는 데이터를 가져오는 방법



Web Scraping with Python

파이썬으로 웹 크롤러 만들기

한빛미디어

라이언 미켈 지음
한민용 옮김

웹 데이터 가져오기

- 웹에서 데이터 가져오기

The screenshot shows a web browser displaying the Naver movie page for '원더 우먼 1984' (Wonder Woman 1984). The browser's developer tools are open, showing the 'Elements' panel on the right. A red box highlights the 'table.list_netizen' element in the DOM tree, which contains a list of user reviews. The 'Styles' panel on the far right shows the default styles for this element. The 'Console' panel at the bottom is also visible, showing 'What's New' messages.

Movie: 원더 우먼 1984
개봉후 평점 ★★★★★ 5.43 / 10 (참여인원 3511명)

현재 상영작 평점보기
[현재 상영작]

table.list_netizen 563 × 639

총 3,490개의 평점이 있습니다.

원더 우먼 1984 ★★★★★ 2	1984년도 CG를 제대로 감상할 수 있습니다! 신고	21.01.04
원더 우먼 1984 ★★★★★ 1	1점 아까운 용가리 이후 최악의 영화 신고	21.01.04
원더 우먼 1984 ★★★★★ 2	1984년에 개봉한 영화 보느라 신고	21.01.04
원더 우먼 1984 ★★★★★ 10	좋아요 ㅎㅎ 좋아요 ㅎㅎ 좋아요 ㅎㅎ 신고	21.01.04
원더 우먼 1984 ★★★★★ 1	첫오프닝 할인쿠폰 같은 장면이 좋았... 너무 지루하고 어린이 만화 실사판 보 는 것 같았다... cg도 어색하고, 틀림해서 연기한 배우들이 대단한듯... 신고	21.01.04
원더 우먼 1984 ★★★★★ 1	감독님이 속제를 급조로 하셨네요. DC역사에 남을 만한 활약을 내놔야 합니다 여 러분... 신고	21.01.04
원더 우먼 1984 ★★★★★ 1	우뢰매가 될 닷다.. 신고	21.01.04
원더 우먼 1984 ★★★★★ 6	원더 우먼 1984 ★★★★★ 6	1984년...

- 웹에서 데이터 가져오기

[illegible]

RStudio

