
Introduction to R packages

Contents

1.

About R Packages

2.

R Download & Installation

3.

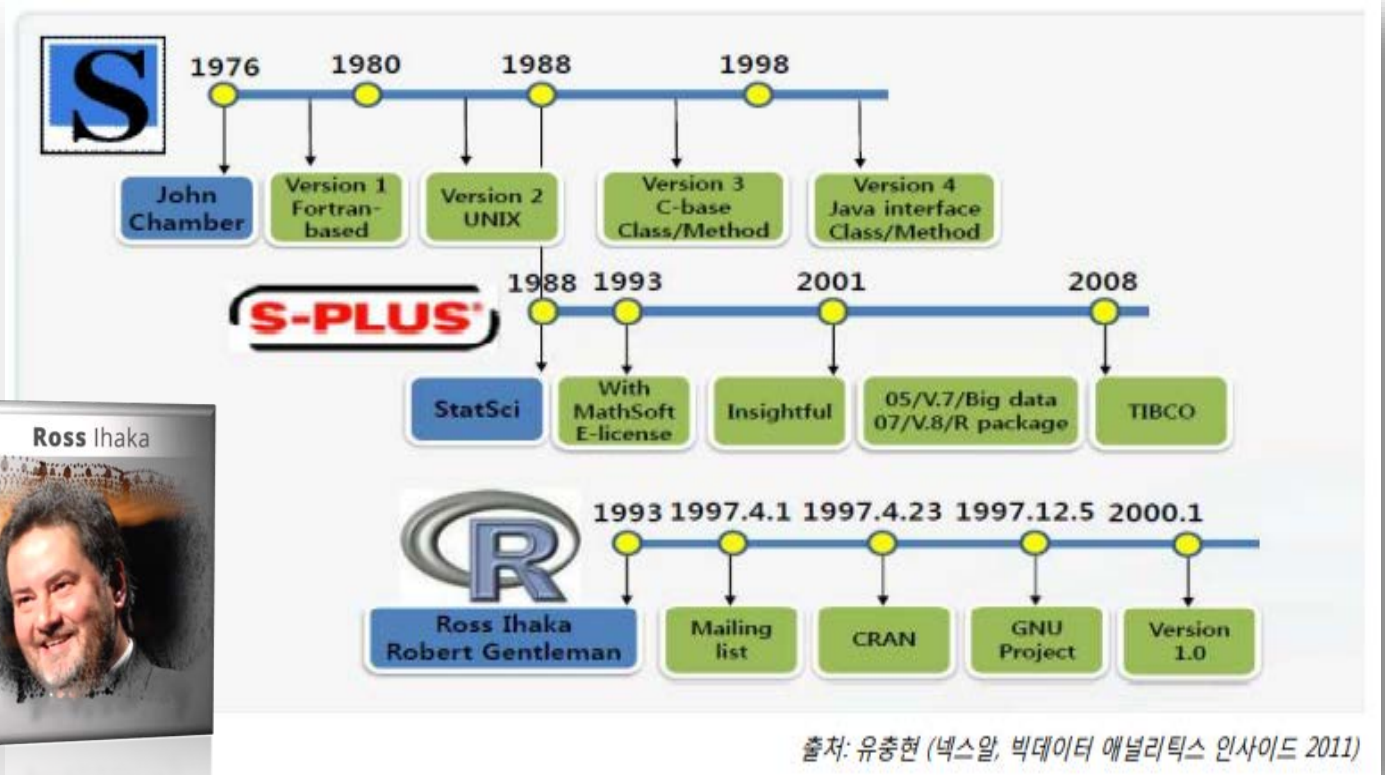
R-Programming

4.

R Graphics

1. About R packages

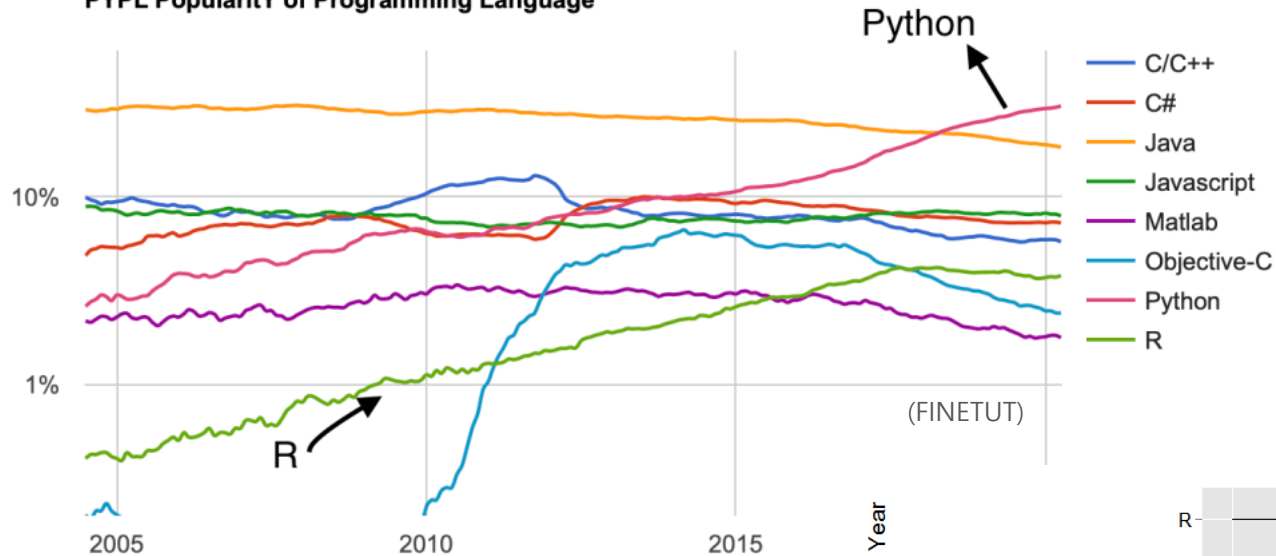
Founders of R



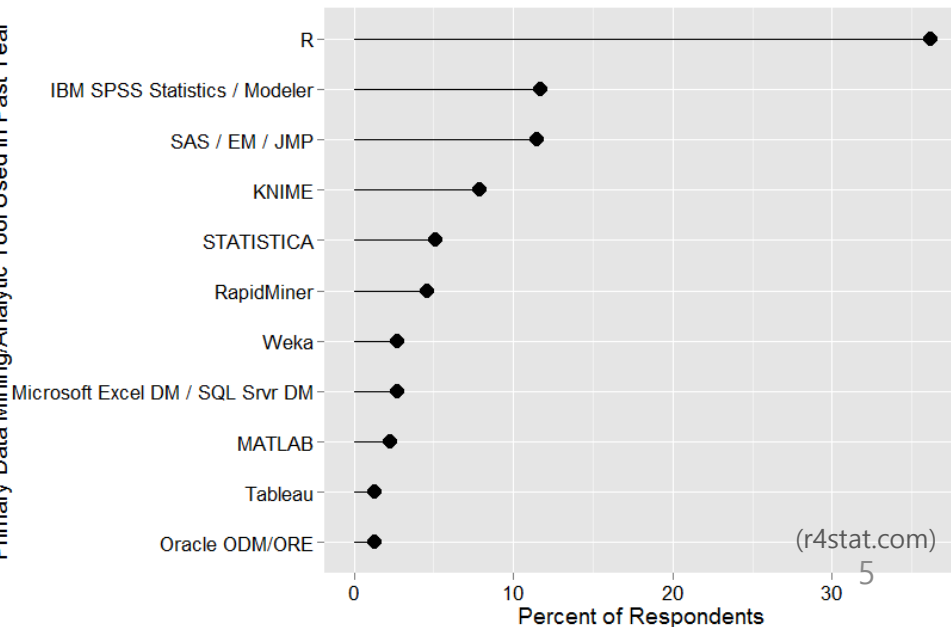
- **R 패키지**는 1990년대 중반 **Robert Gentleman** (23andMe)과 **Ross Ihaka** (Univ. of Auckland)에 의해 free software로 처음 개발되었다.
- 1976년 Bell Lab에서 개발된 **S Language**가 전신이다.

Popularity of R

PYPL Popularity of Programming Language



Primary Data Mining/Analytic Tool Used in Past Year



(r4stat.com)

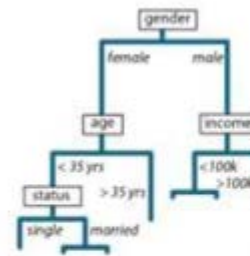
기업체에서의 R의 활용

- 빅데이터 (Big Data)를 활용하는 관련기업의 분석 플랫폼 엔진으로 사용 중
- 특히, Google과 Facebook은 R을 자사의 주된 분석 플랫폼으로 활용



How Google and Facebook are using R

by mike | February 19th, 2009



(March 26th Update: Video now available)

Last night, I moderated our Bay Area R Users Group kick-off event with a panel discussion entitled "The R and Science of Predictive Analytics", co-located with the Predictive Analytics World conference here in SF.

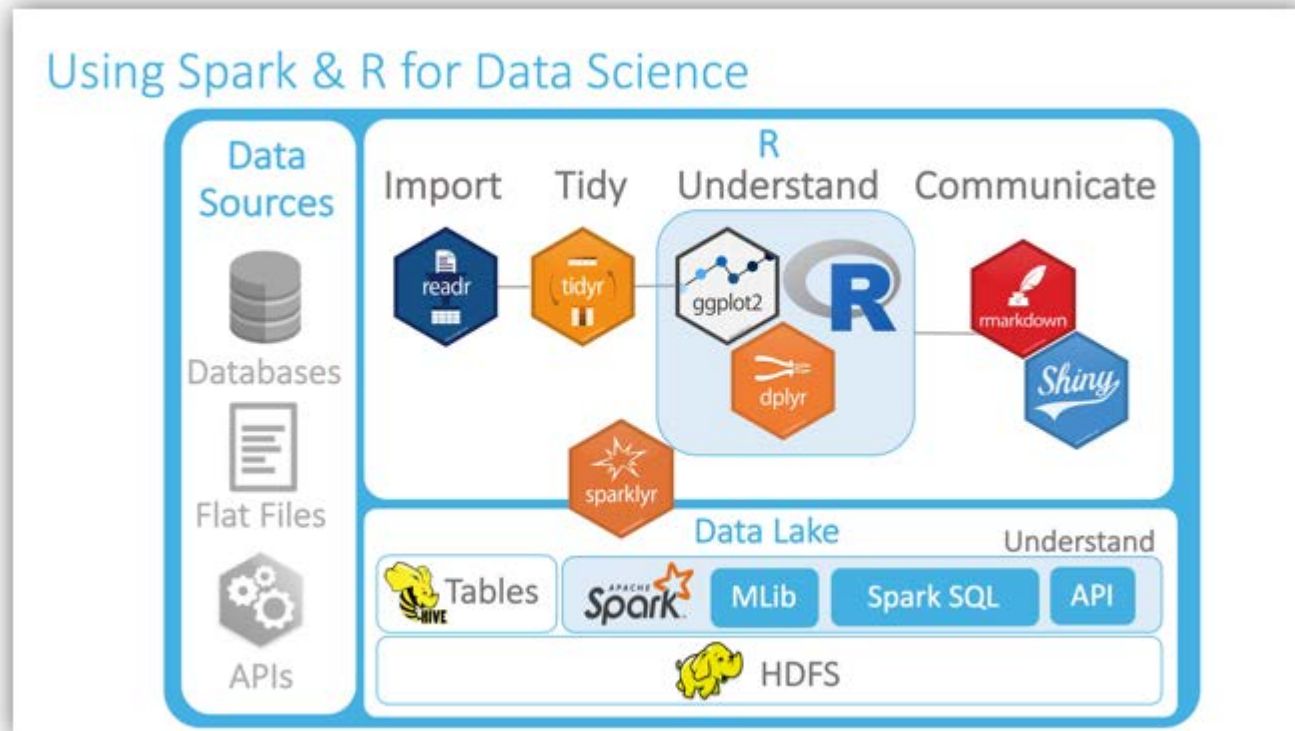
The panel comprised of four recognized R users from industry:

Bo Cowgill, Google
Itamar Rosenn, Facebook
David Smith, Revolution Computing
Jim Porzak, The Generations Network (and Co-Chair of our R Users Group)

The panelists were asked to explain how they use R for predictive analytics within their firms, its strengths and weaknesses as a tool, and provide a case study. What follows is my summary with comments.

빅데이터 분석용 소프트웨어

- 빅데이터 분석을 위한 아키텍처 전반에 걸쳐 공통적인 분석 플랫폼으로 자리 잡음



R through Excel

- 빅데이터 뿐 아니라 통계 비전문가들을 위한 플랫폼도 다양하게 개발됨



About R

R is a language and environment for statistical computing and graphics.

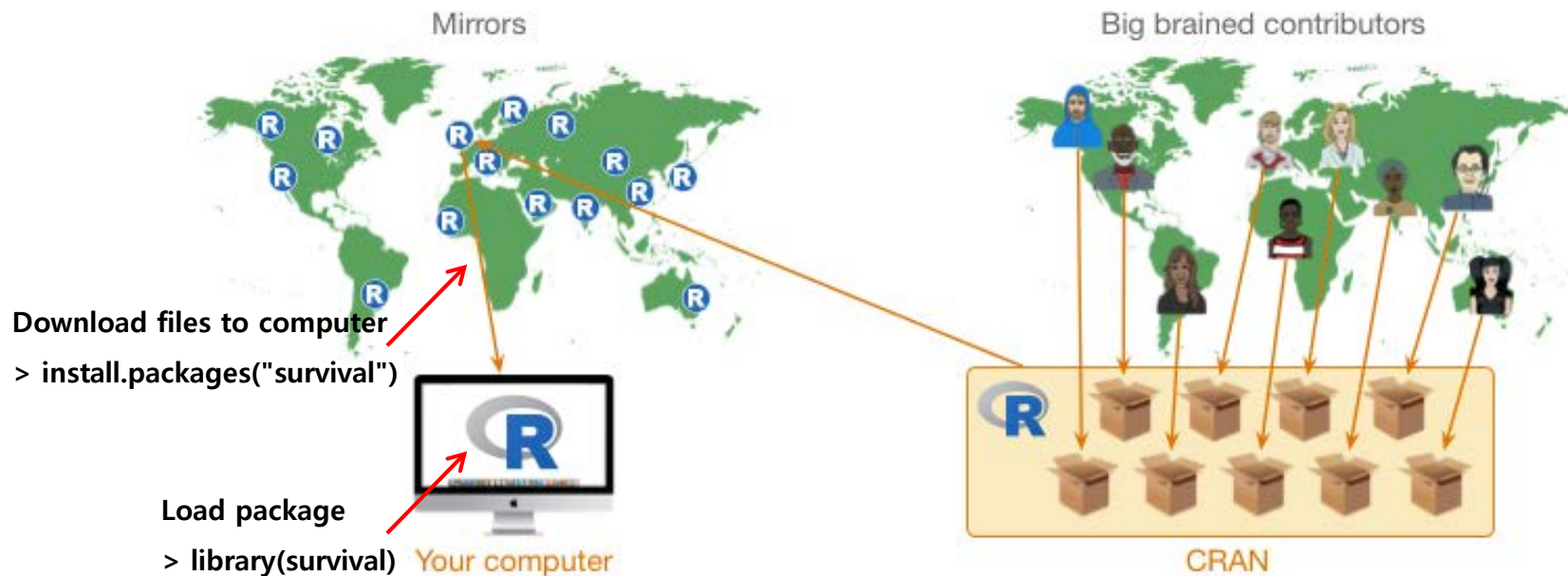
- R은 통계 계산을 위한 시스템으로 데이터 분석 및 그래프 작성을 위한 환경을 지닌다.
- R은 free open source 이다.
 - 언제 어디서든 다운로드 및 설치가 가능
 - Windows, Linux, Unix, Mac OS 등 다양한 운영체제에서 동작
 - 누구나 패키지를 만들어 다른 유저들과 공유 가능
 - Java, Python, Visual Studio 등 다양한 개발 언어 및 플랫폼과 연동
- 데이터 입출력, 데이터 처리, 데이터 분석, 그래프 작성 등을 위한 수많은 알고리즘 및 방법론을 제공한다.

Characteristics of R

- In-Memory Computing
 - 장점: 빠른 처리 속도
 - 단점: H/W 메모리 크기에 영향을 받음
- Object-oriented programming
 - 데이터와 함수가 object로 관리
- Packages
 - 최신의 알고리즘 및 방법론을 적용
 - 다양한 함수 및 데이터 내장
 - Help 창에서 Examples을 바로 사용 가능

R Packages

- R은 패키지 라는 구성요소를 가지고 있으며, 패키지 안에는 여러 함수 형태로 된 코드들이 담겨 있다. 패키지는 특정 기업/기관에서 주관하지 않고 사용하는 모든 user들로부터 개발되어지고 있다.



(Milton Rocks)

R CRAN

What is the CRAN mirror?

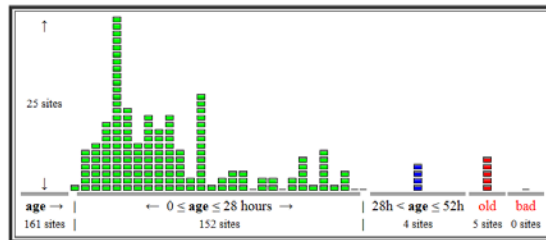
- The **C**omprehensive **R** **A**rchive **N**etwork
- 접근하기 쉽고 주기적으로 자주 업데이트되는 CRAN 선택
- https://cran.r-project.org/mirmon_report_release.html
- http는 안전하게 전송할 수 있고, https 빠르게 전송 가능하다.

Currently, the CRAN package repository features 12503 available packages.
 Currently, the CRAN package repository features 13697 available packages.
 Currently, the CRAN package repository features 14972 available packages.
 Currently, the CRAN package repository features 16214 available packages.

the status of [CRAN/bin/windows/contrib/r-release](https://cran.r-project.org/mirmon_report_release.html) mirrors

date : Mon May 14 08:08:20 2018 (UTC)
 last check : Mon May 14 08:08:20 2018 (UTC)

age histogram



regions

[ar](#) [asia](#) [at](#) [au](#) [be](#) [bg](#) [br](#) [ca](#) [ch](#) [cl](#) [cn](#) [co](#) [cz](#) [de](#) [dk](#) [dz](#) [ec](#) [ee](#) [es](#) [fr](#) [gr](#) [hu](#) [id](#) [ie](#) [in](#) [ir](#) [is](#)

report

161 sites in 48 regions				
0 bad -- 5 older than 2.2 days -- 4 unreachable for more than 5 hours				
last probes : 157 were ok, 4 had no time				
mean mirror age is 21 hours, std_dev 3.3 days, median 8 hours				
CRAN/bin/windows/contrib/r-release site -- home	type	mirror age, daily stats	last probe, probe stats	last stat
Italy				
cran_mirror.garr.it @	http	4 hours	3 hours	ok
dssm.unipa.it @	http	16 hours	2 hours	ok
cran.stat.unipd.it @	http	10 hours	3 hours	ok
cran.stat.unipd.it @	https	10 hours	3 hours	ok
Japan				
cran.ism.ac.jp @	http	1 hour	renewed	ok
cran.ism.ac.jp @	https	3 hours	2 hours	ok
ftp.yz.yamagata-u.ac.jp @	https	6 hours	2 hours	ok
Korea, Republic of				
healthstat.snu.ac.kr @	http	25 hours	2 hours	ok
cran.biodisk.org @	http	19 hours	1 hour	ok
cran.biodisk.org @	https	19 hours	renewed	ok

R packages vs. other programs

	SPSS	SAS	R
Current Version (Sep. 2020)	26.0	9.4	4.0.2
Creator	Norman H. Nie, Dale H. Bent, and Hadlai (Tex) Hull	Jim Goodnight and Jim Barr, North Carolina State University	Ross Ihaka and Robert Gentleman, University of Auckland, New Zealand and the R foundation
Year of Released	1968	Mass distributed since 1972	1995
User Interface	Point-and-click	Based code	Based code
Cost	Expensive (monthly)	Expensive (yearly)	Open source
Advantage	Simple to handle Similarity with Excel	Large scale data	Free

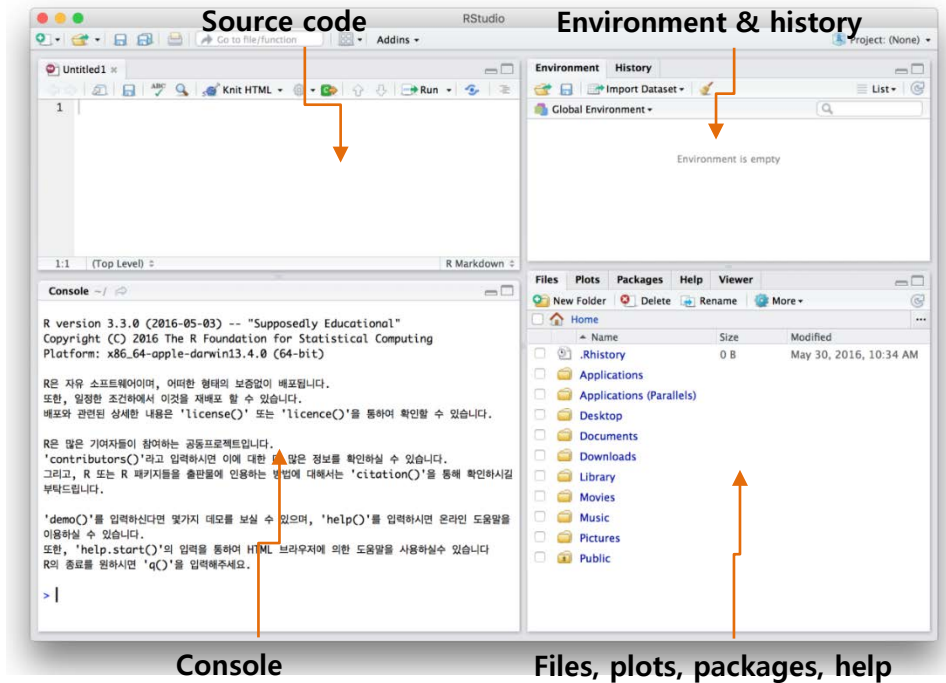
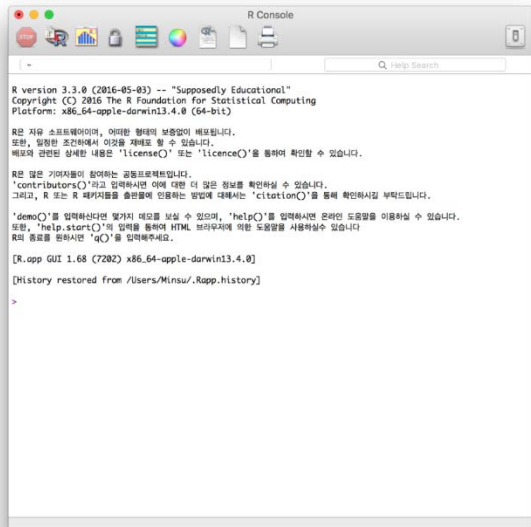
2. R Download & Installation

Download of R

- R과 R studio 다운로드는 아래 주소에서 가능하다.

- R base: <http://cran.r-project.org/>

- R studio: <http://www.rstudio.com/>



Installation of R (www.r-project.org)

The R Project for Statistical Computing

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of platforms.

Download CRAN

Download R for Windows

base

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, windows: both [graphical](#) and [command line versions](#) are available.

Information and Database Systems Laboratory, Pukyong National University
Yeungnam University
Bigdata Campus, Seoul Metropolitan Government
Graduate School of Public Health, Seoul National University, Seoul
The Genome Institute of UNIST (Ulsan National Institute of Science and Technology)
The Genome Institute of UNIST (Ulsan National Institute of Science and Technology)

Installation of R studio

(<https://www.rstudio.com/products/rstudio/download/>)



[Products](#) [Resources](#) [Pricing](#) [About Us](#) [Blogs](#) [Q](#)

Choose Your Version of RStudio

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.](#)



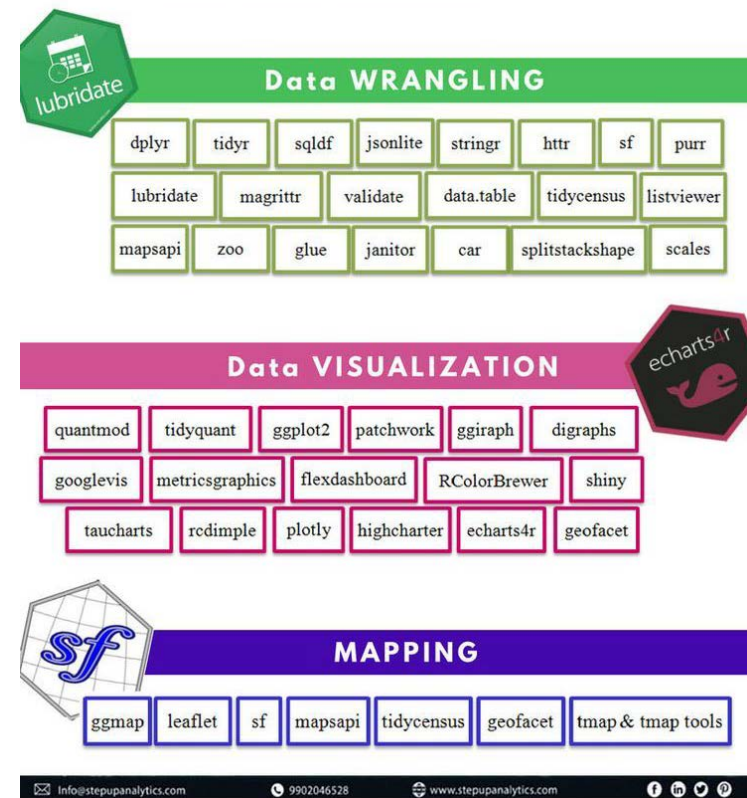
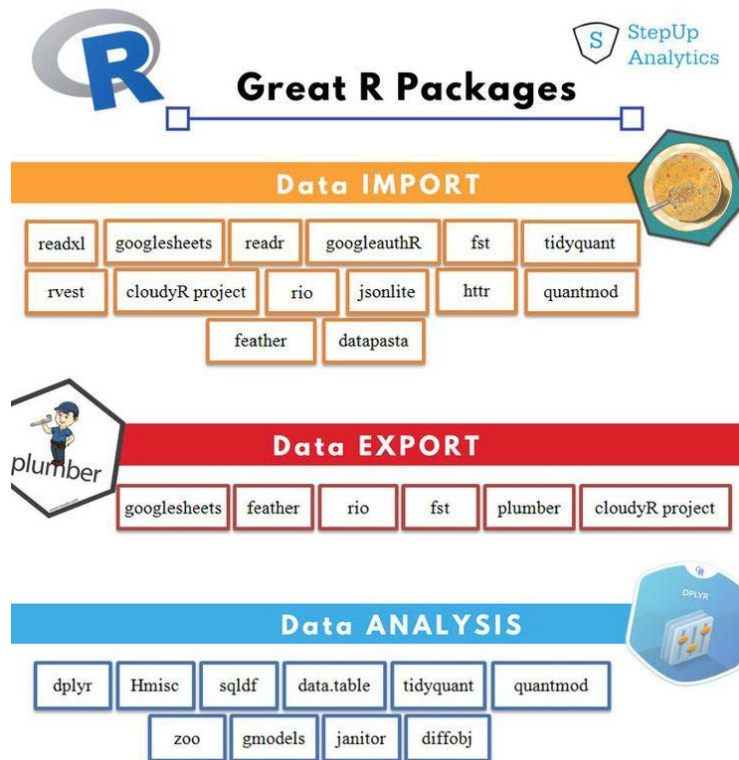
RStudio Desktop Open Source License	RStudio Desktop Commercial License	RStudio Server Open Source License	RStudio Server Pro Commercial License	RStudio Server Pro + RStudio Connect Commercial License
FREE	\$995 per year	FREE	\$9,995 per year	\$29,995 per year
DOWNLOAD Learn More	BUY Learn More	DOWNLOAD Learn More	DOWNLOAD Learn More	TALK Learn More
Integrated Tools for R				

Installers for Supported Platforms

Installers	Size	Date	MD5
RStudio 1.1.463 - Windows Vista/7/8/10	85.8 MB	2018-10-29	58b3d796d8cf96fb8580c62f46ab64d4
RStudio 1.1.463 - Mac OS X 10.6+ (64-bit)	74.5 MB	2018-10-29	a79032ba4d7daaa86a8da01948278d94
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	89.3 MB	2018-10-29	8a6755fa9fae2bafce289df3358aaf63

R Packages

- 목적에 따라 다음의 R 패키지들을 활용할 수 있습니다.



Info@stepupanalytics.com

9902046528

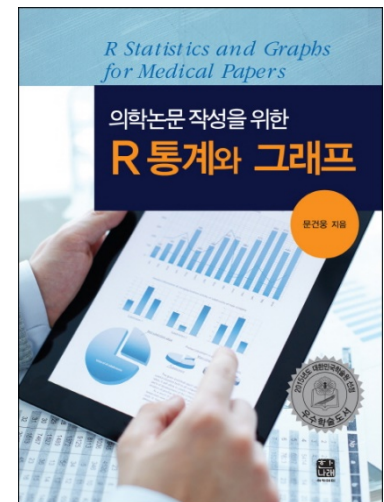
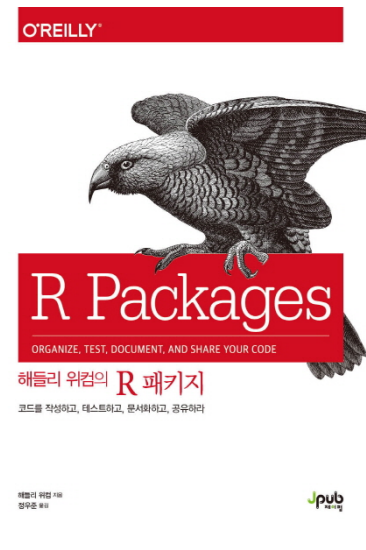
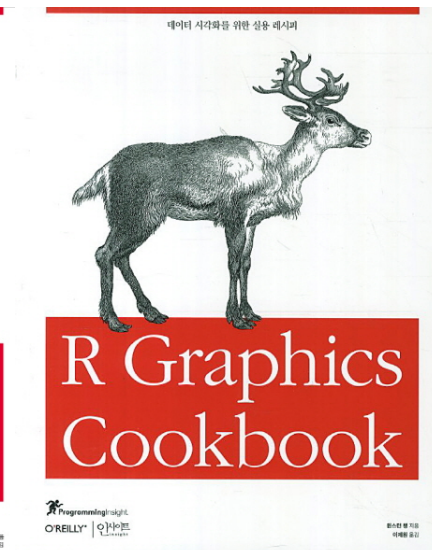
www.stepupanalytics.com

Facebook, Twitter, LinkedIn, YouTube icons

(StepUp Analytics)

Books for R

- R Cookbook (Paul Teetor, 이제원 옮김)
- R Graphics Cookbook (Winston Chang, 이제원 옮김)
- 해들리 위컴의 R Packages (Hadley Wickham, 정우준 옮김)
- 의학논문작성을 위한 R통계와 그래프 (문건웅)



References for R

- <http://www.r-project.org>
- <http://r-project.kr> (한국 R 사용자 모임 – KRUG)
- <http://www.r-bloggers.com>
- <http://stackoverflow.com>
- <http://stats.stackexchange.com>
- <http://www.inside-r.org/>
- <http://www.r-statistics.com/>

Before beginning

- Window 64bit의 경우 "R x64 Version"을 사용하고 32bit의 경우 "R i386"을 사용함
- R Console 창에 직접 입력하는 것보다는 R Editor를 사용하는 것을 권장
- R Editor에서 코드 실행
 - R base: Ctrl+R or F5
 - R Studio: Ctrl+Enter (적용 후 cursor는 다음 line이동)
Alt+Enter (적용 후 cursor는 제자리)
- R Editor를 닫으면 기록된 code가 모두 삭제되므로 자주 저장을 해야 함
- R에서는 대소문자를 구분하기 때문에 주의해야 함 (예. Age, AGE, age는 모두 다른 변수)
- "#" 는 주석(annotation)을 작성하기 위한 시작 코드이며, 같은 line에서만 적용됨
- 다음 페이지부터 모든 코드 앞에는 ">" 표시가 있으며 해당 기호를 붙여 실행하면 오류 발생
- 슬라이드에 있는 " "와 R Editor에서 입력되는 " "가 다르기에 그대로 붙이면 오류 발생

3. R Programming

Thank you!
