

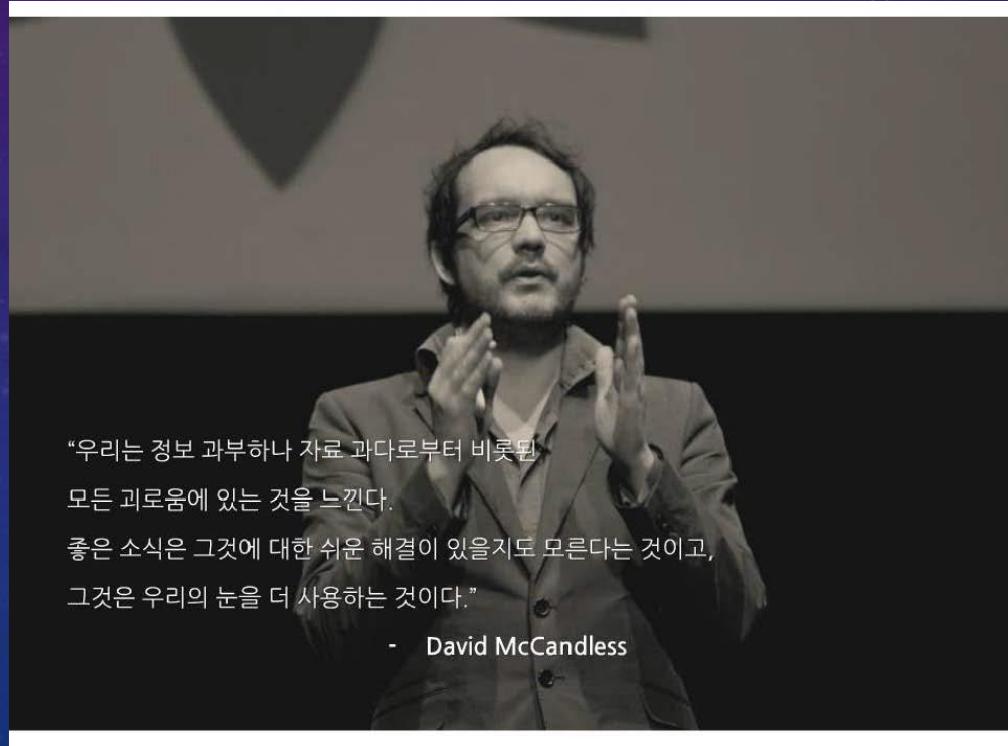
R 데이터 시각화

통계교육원 (2018.8.27~2018.8.29)

대구대학교 수리빅데이터학부 윤상후

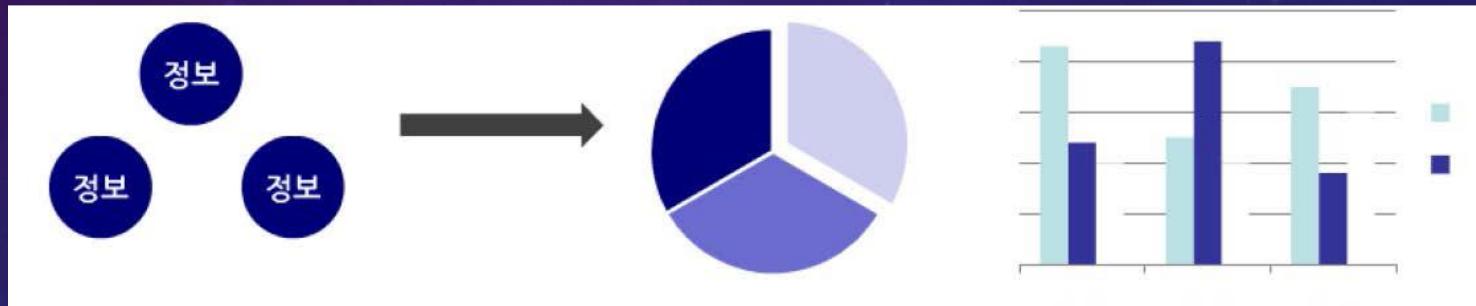
빅데이터와 데이터 시각화

- R을 이용한 데이터 시각화를 소개하고 이를 통해 시각화 이해



데이터시각화란?

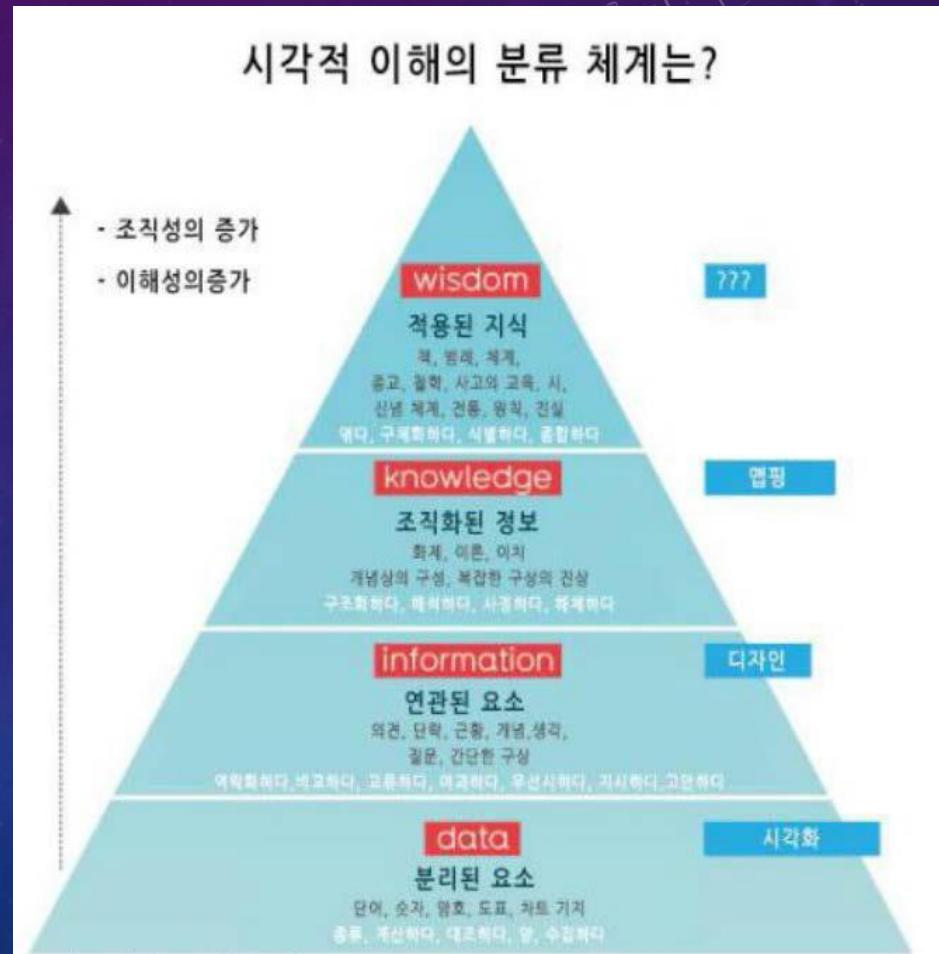
- 광범위하게 분산된 방대한 양의 자료를 분석해 한눈에 볼 수 있도록 도표나 차트 등으로 정리하는 것



자료로부터 정보를 습득하는 시간 절감과 즉각적인 상황 판단 가능
자료를 습득하는 사람들의 흥미 유발 및 정보의 빠른 확산 촉진
자료를 기억하는데 도움

시각 이해의 위계

정보를 전달하여 이해시킨다는
측면에서 각각의 이해를 구축하고
있는 요소



맥캔들레스가 그린 시각적 이해의 위계

<http://www.informationisbeautiful.net/2010/data-information-knowledge-wisdom>

R 을 이용한 데이터 시각화 사례

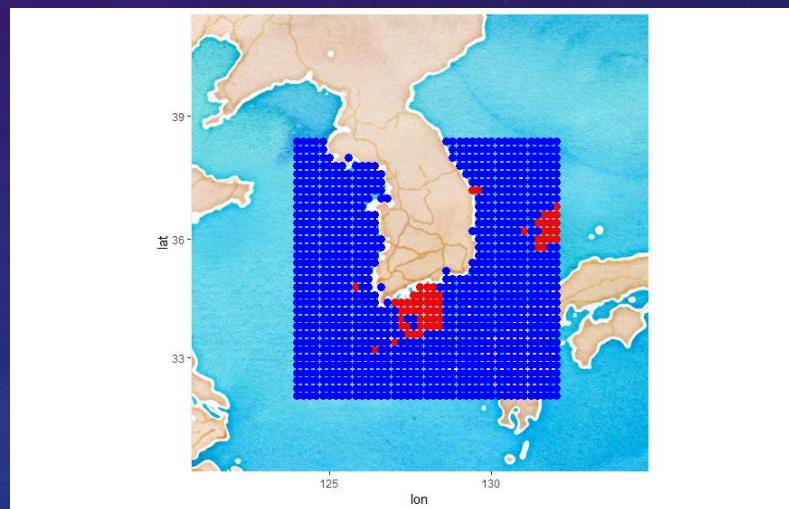
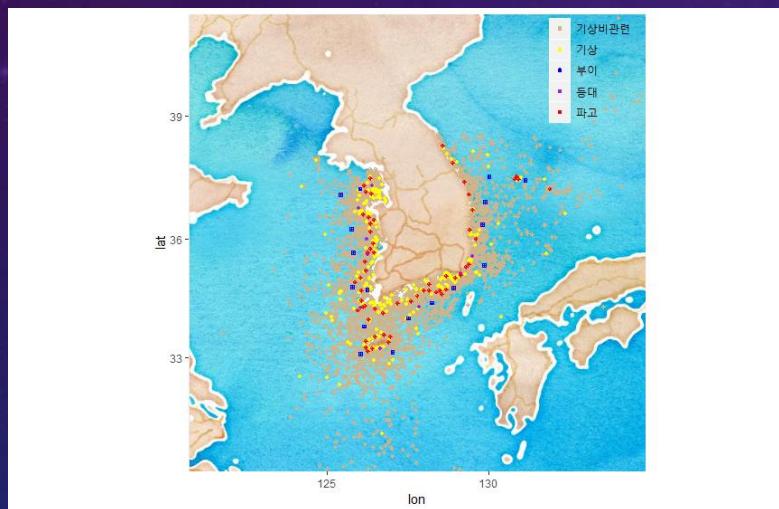
R 시각화 사례와 코드 제공 <https://www.r-graph-gallery.com/>

The screenshot shows the homepage of The R Graph Gallery. At the top, there's a navigation bar with links for HOME, GGPLOT2, ALL GRAPHS, BLOG, ABOUT, PYTHON, and DATA TO VIZ. Below the navigation is a large banner with a dark background and purple circular data points, featuring the text "ART FROM DATA". To the left of the banner is the website's logo, "THE R GRAPH GALLERY", with a stylized orange heart icon. The main content area contains several sections: "Welcome to the R Graph Gallery", "Sponsors" (listing books like "Beginner's Guide to Spatial, Temporal and Spatial-Temporal Ecological Data Analysis with R-INLA" and "The fastest way to learn R! R Courses for Professionals"), and a "TRAINING: R, SCALA, STAN" section featuring a DataCamp course by Hadley Wickham.

데이터 시각화 사례 1

- 해양수산부 공공 빅데이터 활용 경진대회(우수상)

실시간 바다기상자료를 수집하고, 해양사고 피해사례와 연관성을 파악하여
현재 해상의 위험성을 시각화



데이터 시각화 사례 2

- 실시간 지진위험지도

기상청에서 과학도 지지자로 느껴하여 학바도 지지의도 새서

Earthquake risk of the Korean peninsula

실시간 한반도 지진위험지도 생성 방법

A Method for Creating Real-Time Earthquake Hazard Map in Korean Peninsula

한국방재학회논문집

2018, vol. 18, no. 1, 통권 89호 pp. 193-198 (6 pages)

발행기관 : 한국방재학회

연구분야 : 공학 > 기타공학

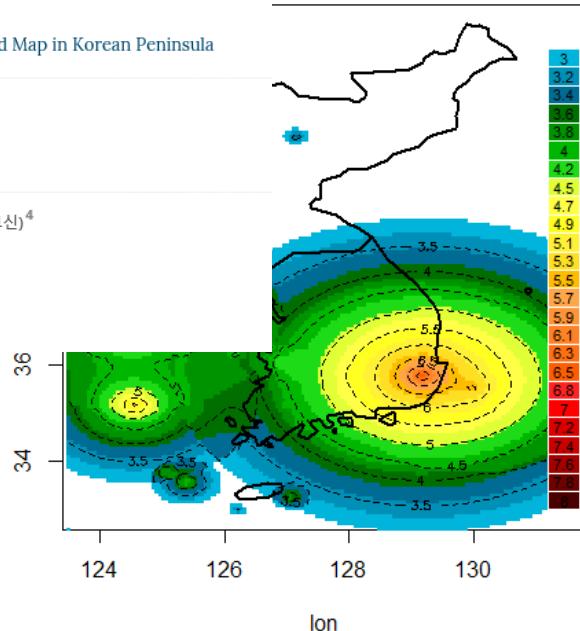
오혜주 (제1)¹, 임지은 (참여)², 이예림 (참여)³, 윤상후 (교신)⁴

¹ 대구대학교

² 대구대학교 통계학과

³ 대구대학교

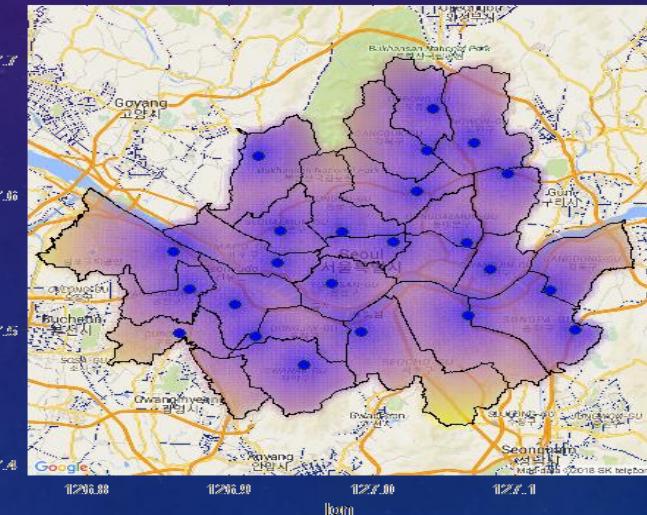
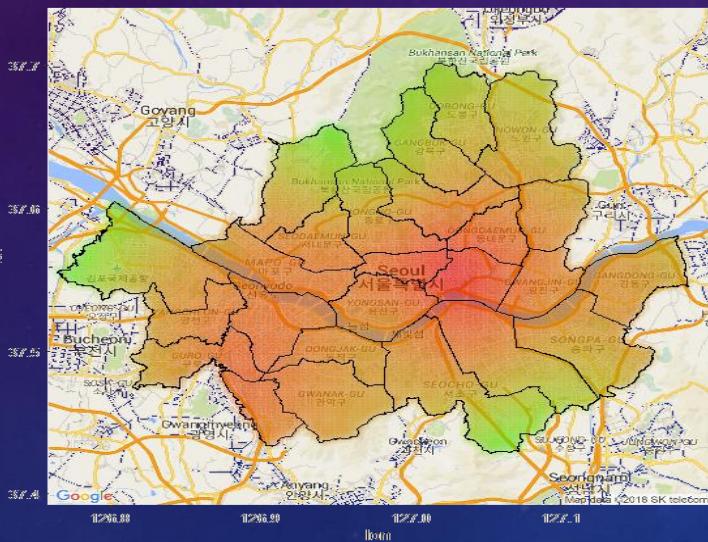
⁴ 대구대학교



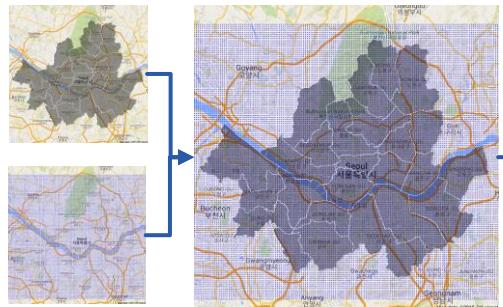
데이터 시각화 사례 3

- 서울특별시 화재 자료 시각화

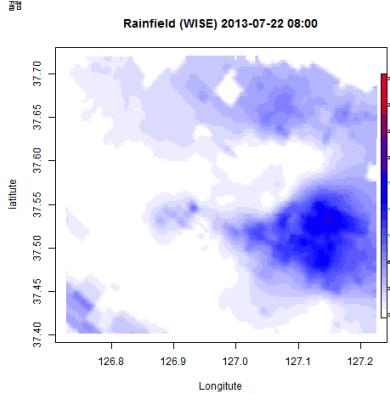
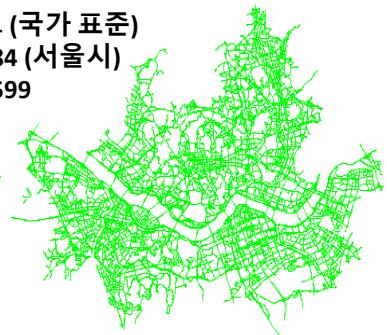
2015~2017년 서울특별시 동별 화재자료와 소방서의 위치를 시각화하여
화재 진압을 위한 소방서의 위치의 적절성 판단



지도 (서울시, 250m 해상도)



표준노드링크 (국가 표준)
링크수 : 22,184 (서울시)
포인트 : 177,599

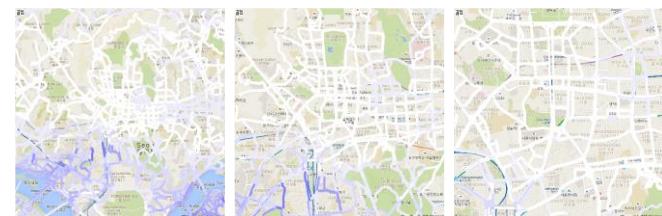
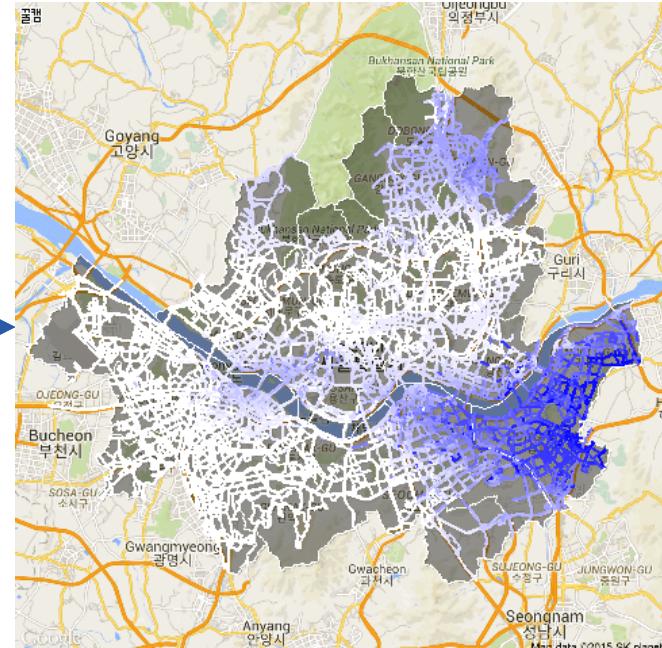
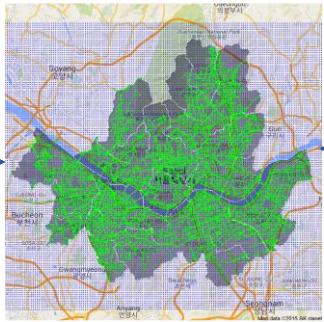


고해상도 실시간 강우장 (250m)

안전주행을 위한 실시간 도로기상 모니터링 시스템

도로기상 모니터링 시스템

도로기상 예측지점



도로정보 & 고해상도 강우장

(250m)

R 이란?

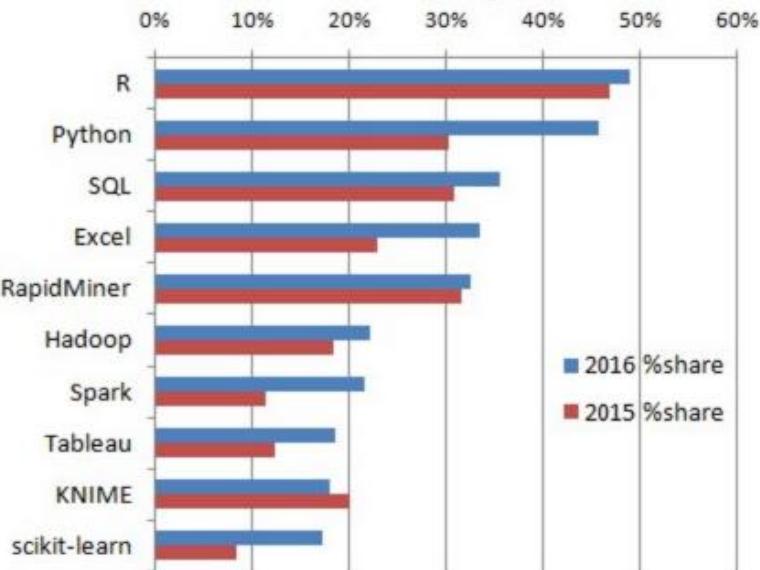
- 통계 계산과 그래픽을 위한 **프로그래밍 언어**
- 뉴질랜드 오클랜드 대학교(Ross Ihaka 교수) 교육용 통계패키지로 개발
- 현재 R 코어 팀에서 “R-project”라는 이름으로 계속 개발중

R 프로젝트 (<http://www.r-project.org>)

- 소프트웨어 다운로드
- 매뉴얼, 참고문헌, 사용자 그룹에 대한 정보

Feb 2017	Feb 2016	Change	Programming Language	Ratings	Change
1	1		Java	16.676%	-4.47%
2	2		C	8.445%	-7.15%
3	3		C++	5.429%	-1.48%
4	4		C#	4.902%	+0.50%
5	5		Python	4.043%	-0.14%
6	6		PHP	3.072%	+0.30%
7	9	▲	JavaScript	2.872%	+0.67%
8	7	▼	Visual Basic .NET	2.824%	+0.37%
9	10	▲	Delphi/Object Pascal	2.479%	+0.32%
10	8	▼	Perl	2.171%	-0.08%
11	11		Ruby	2.153%	+0.10%
12	16	▲	Swift	2.125%	+0.75%
13	13		Assembly language	2.107%	+0.28%
14	38	▲	Go	2.105%	+1.81%
15	17	▲	R	1.922%	+0.73%
16	12	▼	Visual Basic	1.875%	+0.02%
17	18	▲	MATLAB	1.723%	+0.63%
18	19	▲	PL/SQL	1.549%	+0.49%
19	14	▼	Objective-C	1.536%	+0.13%
20	23	▲	Scratch	1.500%	+0.71%

KDnuggets Analytics/Data Science 2016 Software Poll, top 10 tools



R이 많이 사용되는 이유?

통계학 분야의 연구와 교육에서 **표준 계산 언어**

다양한 통계 분석과 우수한 그래픽 방법 제공

- 사용자가 **새로운 함수**를 작성하여 추가 가능
- 다른 사용자가 만든 패키지를 인터넷에서 공개 (**쉽게 공유**)

인터넷을 통해 자유롭게 내려 받아 **무료**로 사용 가능

Window/mac/unix 플랫폼에서 설치 및 운영 가능

The sidebar contains the following links:

- [Home] (Home page)
- Download CRAN (Download R from CRAN)
- R Project
 - About R
 - Logo
 - Contributors
 - What's New?
 - Reporting Bugs
 - Development Site
 - Conferences
 - Search
- R Foundation
 - Foundation Board
 - Members
 - Donors
 - Donate
- Help With R
 - Getting Help

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 UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred [CRAN mirror](#).

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.

News

- [R version 3.5.0 \(Joy in Playing\) prerelease versions](#) will appear starting Friday 2018-03-23. Final release is scheduled for Monday 2018-04-23.
- [R version 3.4.4 \(Someone to Lean On\)](#) has been released on 2018-03-15.
- [useR! 2018](#) (July 10 - 13 in Brisbane) is open for registration at <https://user2018.r-project.org>
- [The R Journal Volume 9/2](#) is available.
- [R version 3.3.3 \(Another Canoe\)](#) has been released on Monday 2017-03-06.
- [useR! 2017](#) took place July 4 - 7 in Brussels <https://user2017.brussels>
- The [R Logo](#) is available for download in high-resolution PNG or SVG formats.

무료다운 (<https://www.r-project.org>)

R 데이터 시각화 강의내용

- 8.27일 (월요일)

2장 R 그래픽스의 기초, **3장** 데이터 시각화 함수

- 8. 28일(화요일)

6장 ggplot2 그래픽스

- 8.29일(수요일)

7장 유용한 플롯, 데이터 시각화 사례