



통계분석 및 그래프 쉽게 그리기

가톨릭대학교 성빈센트병원

문 건 웅

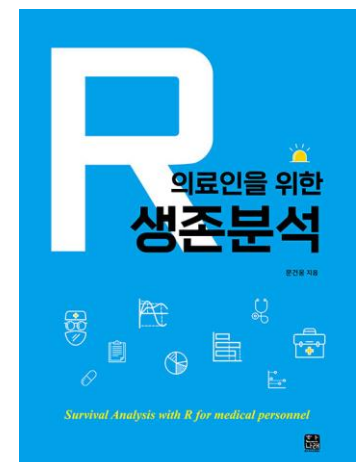
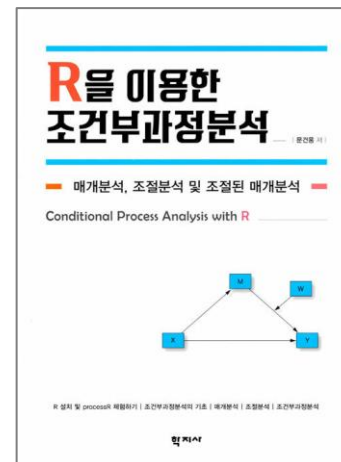
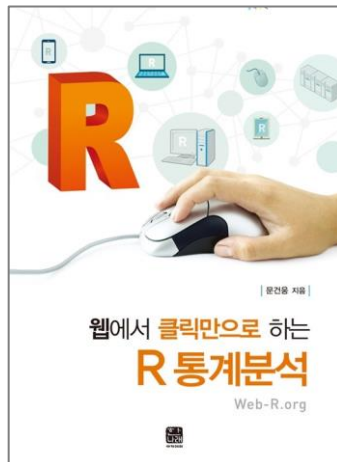
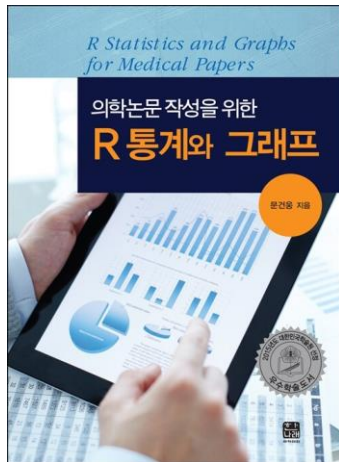
The Korean Society of Cardiology COI Disclosure

Name of First Author: Keon-Woong Moon

The author has no financial conflicts of interest to disclose concerning the presentation

약력

- 가톨릭대학교 성빈센트병원 순환기내과
- 전공 : 중재시술



13 R packages on CRAN

- mycor, moonBook, ztable
- ggiraphExtra, dplyrAssist, editData
- ggplotAssist , webr, rrtable
- predict3d, processR
- autoReg, interpretCI

내용

- Pubmed 논문분석
 - Statistical Software Popularity
 - Programming Languages Popularity
- R과 SPSS, SAS 비교
- R 및 web-R.org 소개

Pubmed Central 논문 분석

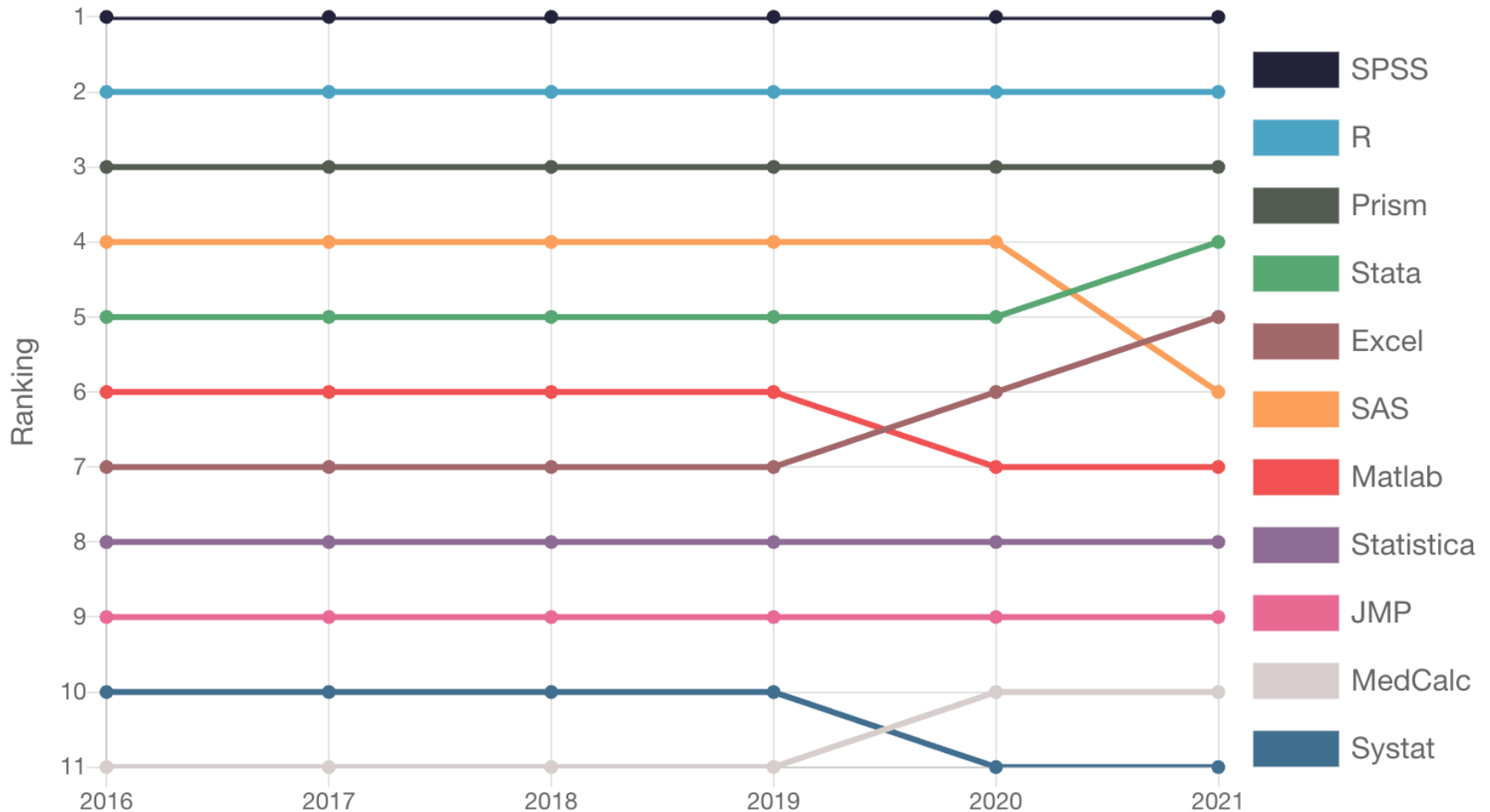
- <https://quantifyinghealth.com/statistical-software-popularity-in-research/>
- George Choueiry





Statistical Software Popularity in 40,582 Research Papers

- 2016 – 2021
- uploaded to PubMed Central
- a random sample of 76,147 full-text research papers
- only 40,582 (53.3%) mentioned the use of at least 1 statistical software.

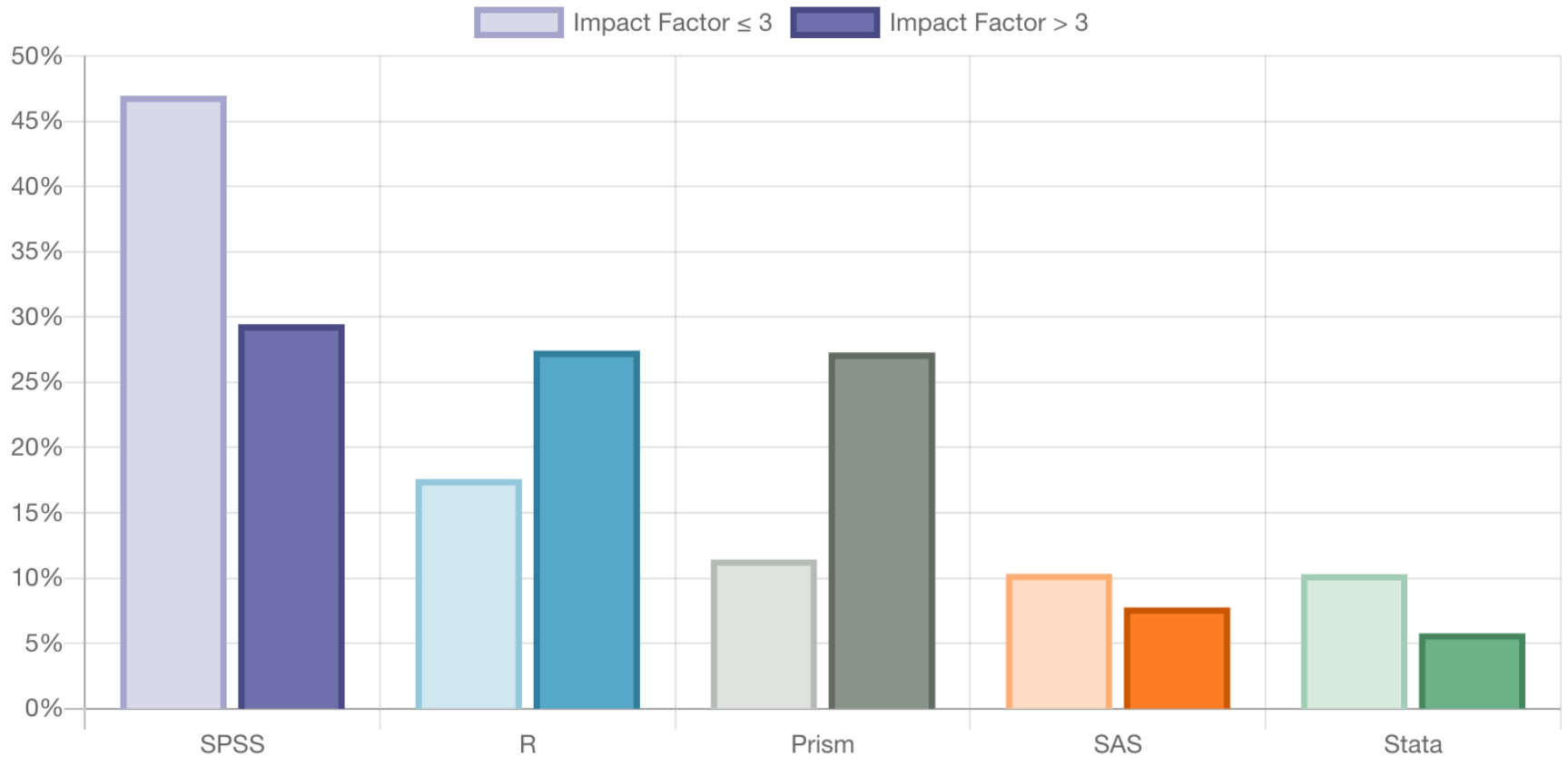
Top statistical packages over the years




Most popular statistical packages overall

Overall Ranking	Software	Number of Mentions (Total: 40,582 articles)	Mentions (in Percent)		6-Year Trend
1	SPSS	16,616	40.48%		-1.43%
2	R	8,423	20.52%		+1.29%
3	Prism	7,132	17.38%		+0.82%
4	SAS	3,693	9.00%		-0.48%
5	Stata	3,386	8.25%		-0.15%
6	Matlab	3,005	7.32%		-0.26%
7	Excel	2,719	6.62%		+0.32%
8	Statistica	763	1.86%		-0.03%
9	JMP	633	1.54%		-0.04%

Low vs High Impact Journals



Programming Languages Popularity in 12,086 Research Papers

Ranking	Programming Language	Number of Mentions (Total: 12086)	Mentions (in Percent)	6-Year Trend
1	R	8423	69.69%	 +1.59%
2	Matlab	2575	21.31%	-2.01%
3	Python	1085	8.98%	+1.21%
4	Java	309	2.56%	-0.24%
5	Perl	301	2.49%	-0.32%
6	C++	151	1.25%	-0.23%
7	JavaScript	97	0.80%	–
8	SQL	94	0.78%	–
9	PHP	87	0.72%	–
10	Visual Basic	40	0.33%	–

목적지 - 통계분석(결과,그림)



원하는 목적지에 가는 방법

- 전용제트기, 운전기사 + 자가용
- 대중교통
 - 택시
 - 지하철, 버스
- 자가운전

전용제트기





전용제트기, 기사

- 의학통계연구소/의학통계전문가 고용
- 장점
 - 운전이 신경쓸 필요가 없다.
- 단점
 - 돈이 아주 많이 든다.
- 하지만
 - 목적지를 정해줘야 한다.

대중교통 - 택시



대중교통 - 택시

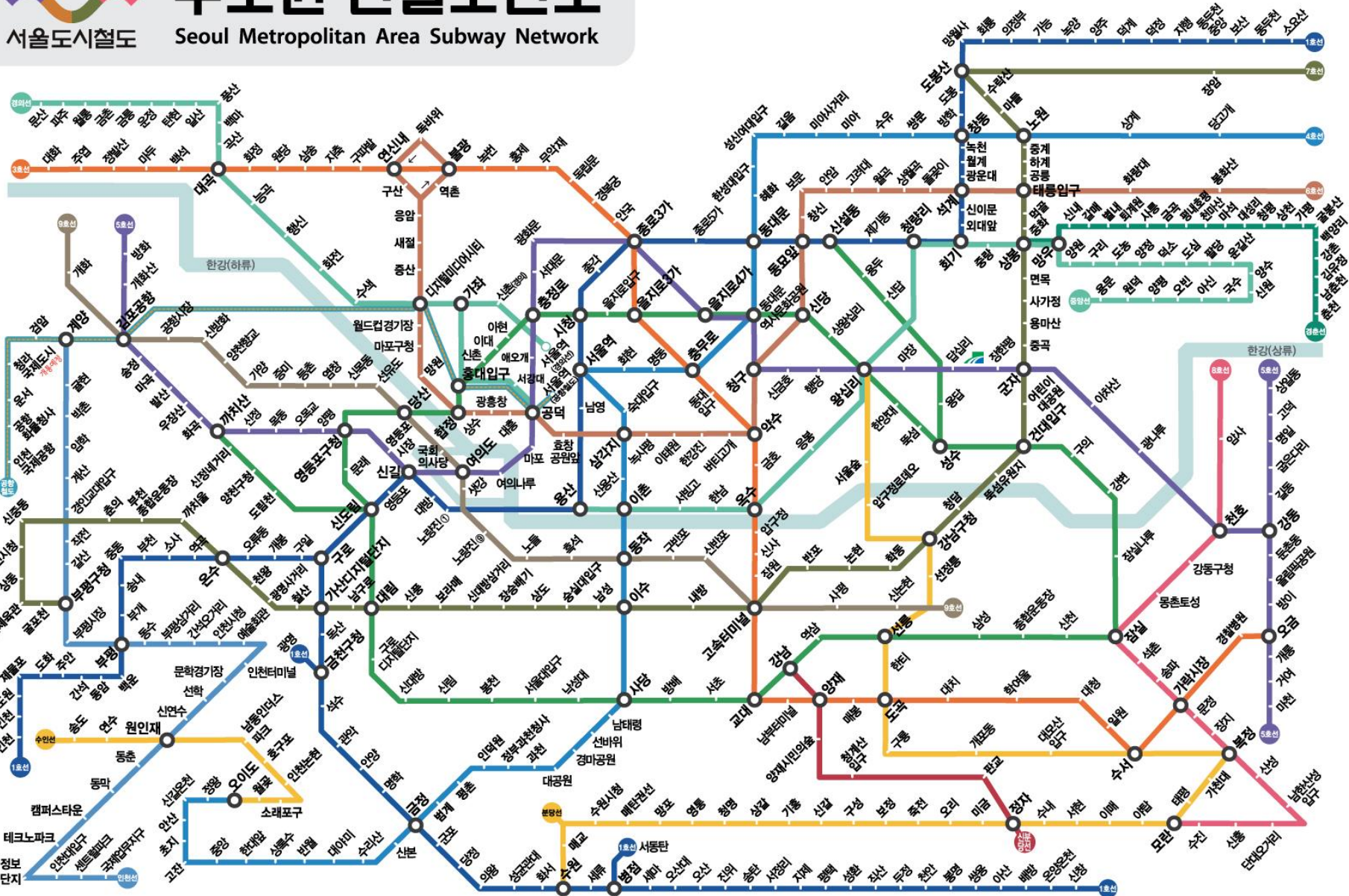
- 통계컨설팅
- 장점
 - 운전을 못해도 된다
- 단점
 - 돈이 든다.
- 특징
 - 목적지를 정해줘야 한다.(연구목적)
 - 데이터에 대한 설명과 데이터 클리닝 필요

대중교통



수도권 전철노선도

Seoul Metropolitan Area Subway Network



대중교통 : 버스, 전철

- SPSS, SAS
- 장점
 - 비용이 비교적 적게 든다.
 - 운전할 필요가 없다
 - 목적지만 알면 갈수 있다
- 단점
 - 정해진 곳만 갈 수 있다.
(정거장,정류장이 없는 곳은 못간다.)

자가운전: R



자가운전 : R

- 단점

- 운전을 배워야 한다(반클러치,기어변속,...)
- 미숙할 경우 시동이 꺼질 수 있다.
- 운전면허가 있어야 한다.
- 도로교통법과 실제 운전 에 익숙해야 한다.
- 네비게이션이 필요한 경우가 많다.

- 장점

- 비용이 가장 적게 든다.
- 멋진 그래프/그림을 그릴 수 있다.
- 아무 곳이나 갈 수 있다(대중교통 없는 곳 포함)
- 남들이 안가는 곳을 갈 수 있다.

New methods



Journal of Statistical Software

July 2022, Volume 103, Issue 3.

doi:10.18637/jss.v103.i03

evgam: An R Package for Generalized Additive Extreme Value Models

Benjamin D. Youngman 
University of Exeter

Abstract

This article introduces the R package **evgam**. The package provides functions for fitting extreme value distributions. These include the generalized extreme value and generalized Pareto distributions. The former can also be fitted through a point process representation. Package **evgam** supports quantile regression via the asymmetric Laplace distribution, which can be useful for estimating high thresholds, sometimes used to discriminate between extreme and non-extreme values. The main addition of package **evgam** is to let extreme value distribution parameters have generalized additive model forms, the smoothness of which can be objectively estimated using Laplace's method. Illustrative examples fitting various distributions with various specifications are given. These include daily precipitation accumulations for part of Colorado, US, used to illustrate spatial models, and daily maximum temperatures for Fort Collins, Colorado, US, used to illustrate temporal models.

Keywords: generalized extreme value distribution, generalized Pareto distribution, point process, generalized additive model, Laplace's method, R.

1. Introduction

Practical extreme value analyses have typically considered modeling block maxima with the generalized extreme value (GEV) distribution or exceedances of a high threshold with the generalized Pareto distribution (GPD); see Davison and Smith (1990) for a seminal work on the latter approach, and Coles (2001) for a detailed overview of both approaches. Here, the GEV and GPD distributions will be considered *the* extreme value distributions (EVD). Smith (1989) develops a model using Pickands' (1971) point process representation of extremes, which, in some sense, marries the two EVDs.

Various packages have been contributed to the Comprehensive R Archive Network (CRAN) to fit EVDs in R (R Core Team 2022). One of the earliest, package **ismev** (Heffernan and Stephenson 2018), allows users to recreate many of the analyses presented in Coles (2001).






Journal of Statistical Software

August 2022, Volume 103, Issue 13.

doi:10.18637/jss.v103.i13

Robust Mediation Analysis: The R Package **robmed**

Andreas Alfons  Erasmus University Rotterdam
Nüfer Y. Ateş  Sabanci University
Patrick J. F. Groenen  Erasmus University Rotterdam

Abstract

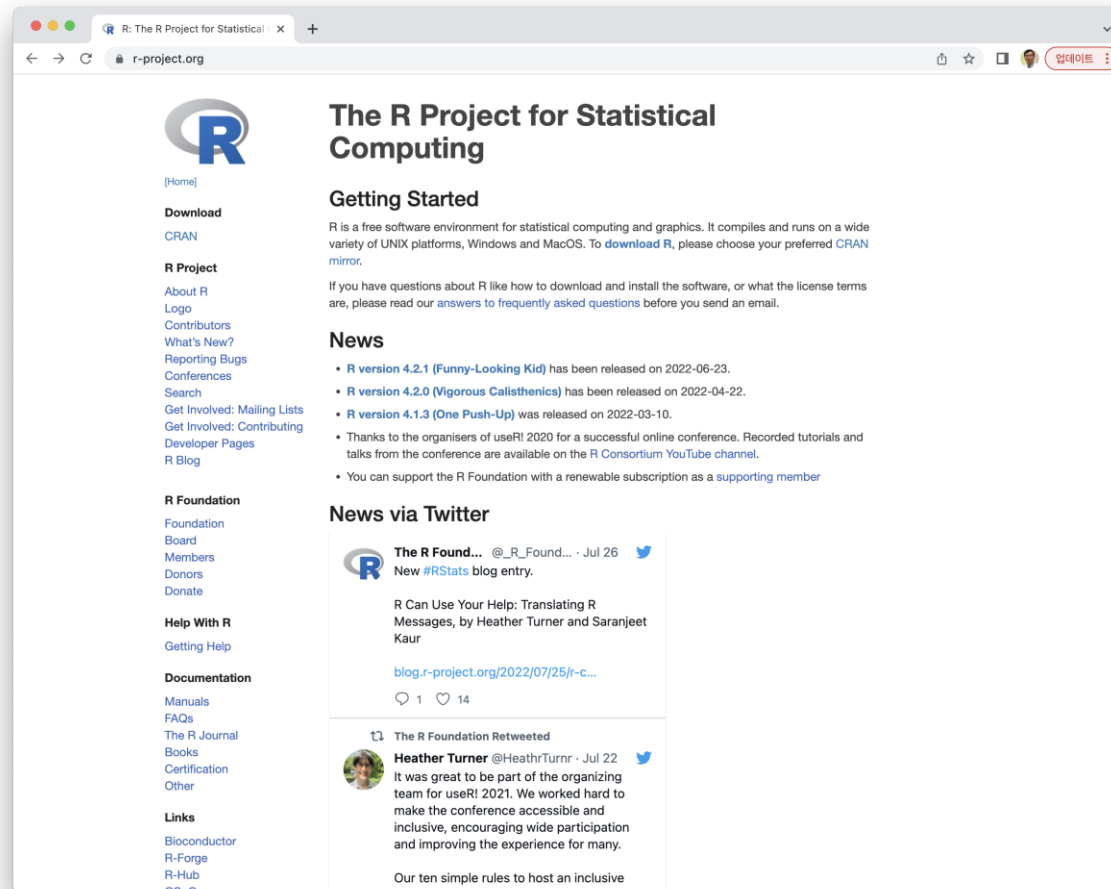
Mediation analysis is one of the most widely used statistical techniques in the social, behavioral, and medical sciences. Mediation models allow to study how an independent variable affects a dependent variable indirectly through one or more intervening variables, which are called mediators. The analysis is often carried out via a series of linear regressions, in which case the indirect effects can be computed as products of coefficients from those regressions. Statistical significance of the indirect effects is typically assessed via a bootstrap test based on ordinary least-squares estimates. However, this test is sensitive to outliers or other deviations from normality assumptions, which poses a serious threat to empirical testing of theory about mediation mechanisms. The R package **robmed** implements a robust procedure for mediation analysis based on the fast-and-robust bootstrap methodology for robust regression estimators, which yields reliable results even when the data deviate from the usual normality assumptions. Various other procedures for mediation analysis are included in package **robmed** as well. Moreover, **robmed** introduces a new formula interface that allows to specify mediation models with a single formula, and provides various plots for diagnostics or visual representation of the results.

Keywords: mediation analysis, robust statistics, bootstrap, R.

1. Introduction

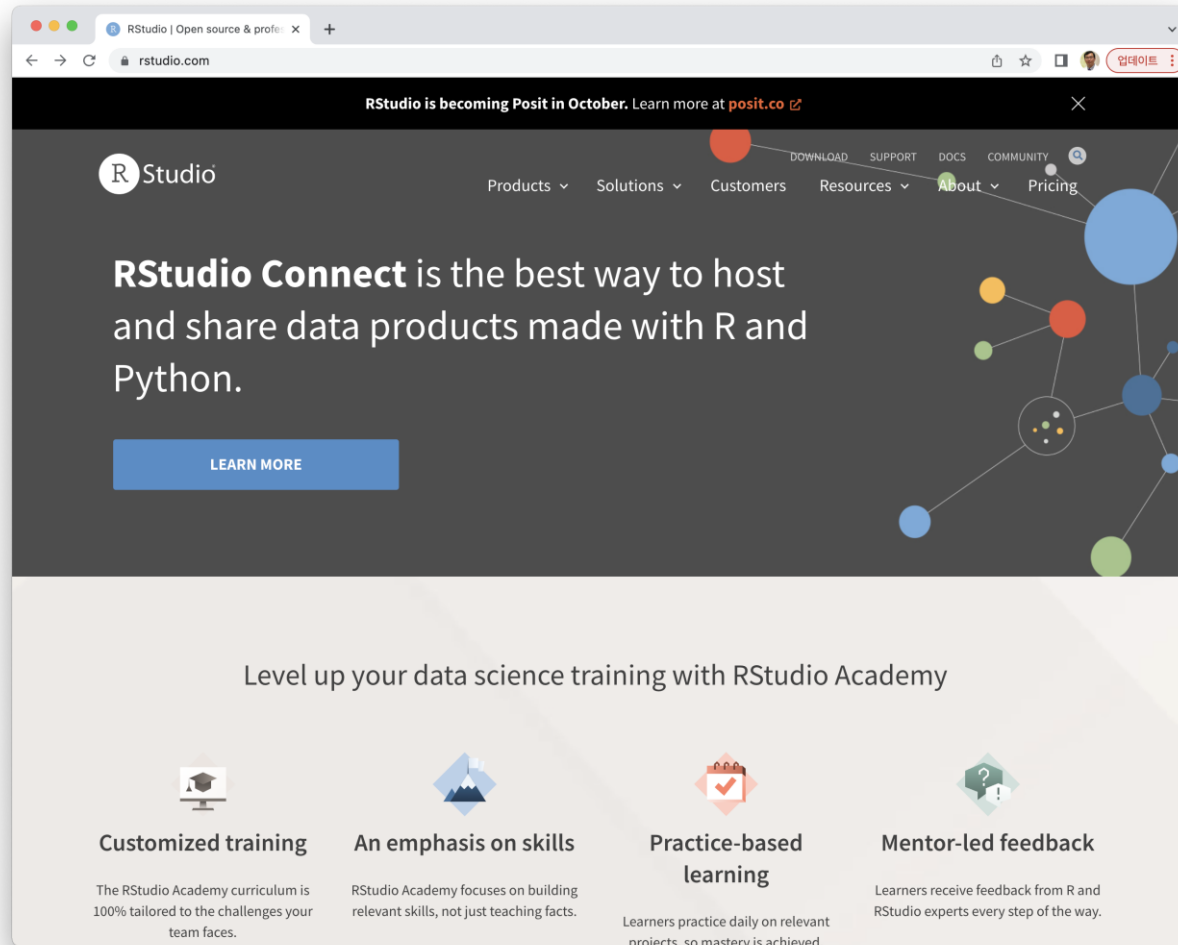
In the social, behavioral, and medical sciences, mediation analysis is a popular statistical technique for studying how an independent variable affects a dependent variable indirectly through an intervening variable called a mediator. For instance, Erreygers, Vandebosch, Vranjes, Baillien, and De Witte (2018) find that poor sleep quality in adolescents explains cyberbullying through anger, and Gaudiano, Herbert, and Hayes (2010) report that the believability of hallucinations after treatment for psychotic disorders mediates the relationship between the type of treatment and distress after treatment. Figure 1 shows a diagram of the

R 설치 r-project.org



RStudio 설치

rstudio.com

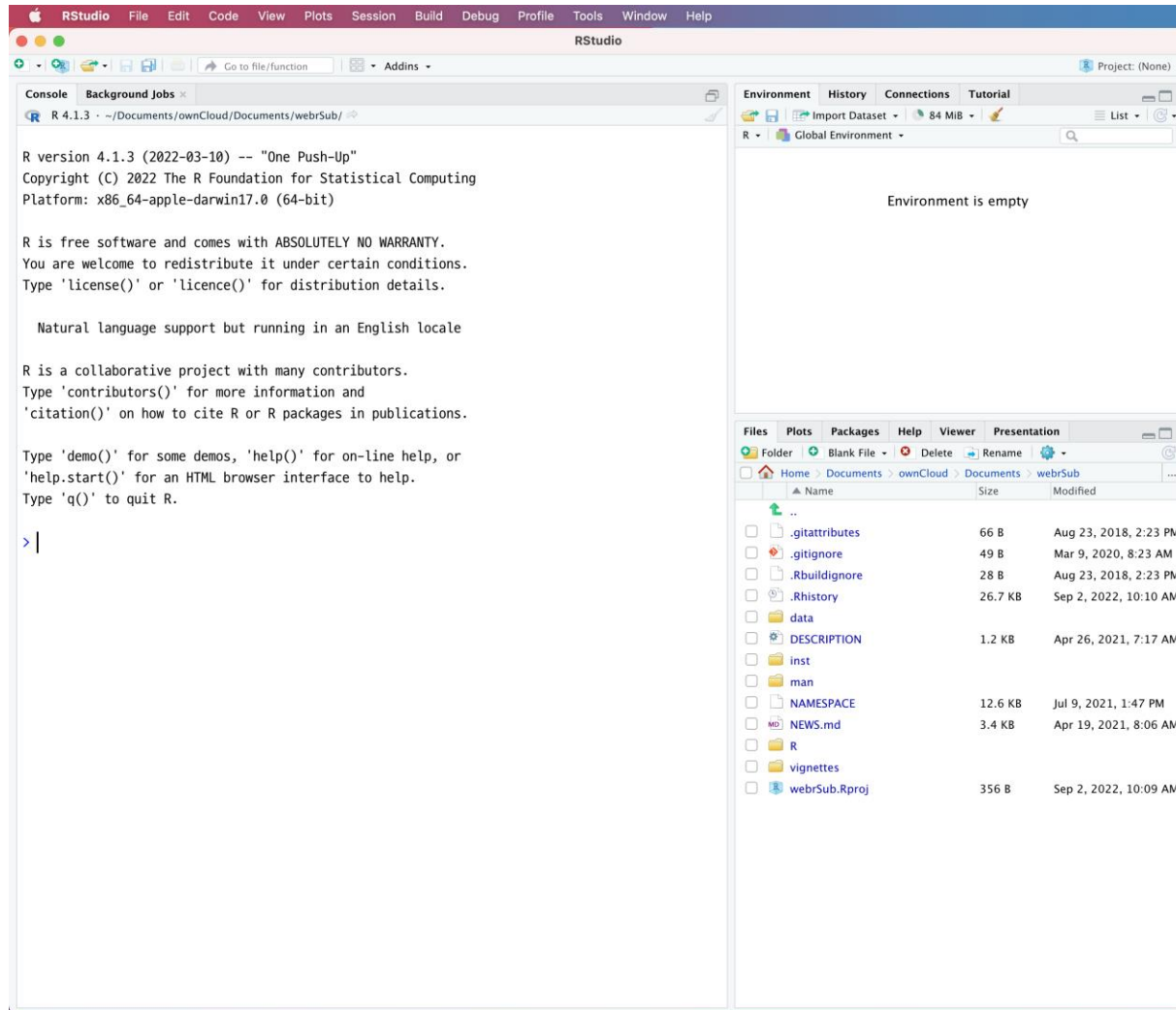


2022.9.23 Fri. - 25 Sun.
EXCO, Daegu

The 66th Annual Scientific Meeting of The Korean Society of Cardiology



RStudio 실행





R 통계분석 나도 할 수 있을까?

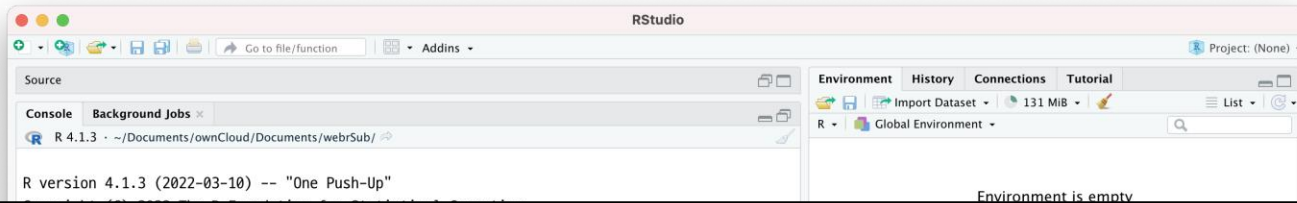


R 통계분석 쉽게(?) 배우기





에러 메시지



```
> library(moonBook)
```

```
> library(MoonBook)
```

```
Error in library(MoonBook) : there is no package called 'MoonBook'
```

```
Backtrace:
```

```
1. base::library(MoonBook)
```

```
>
```

```
> library(moonBook)
```

```
> library(MoonBook)
```

```
Error in library(MoonBook) : there is no package called 'MoonBook'
```

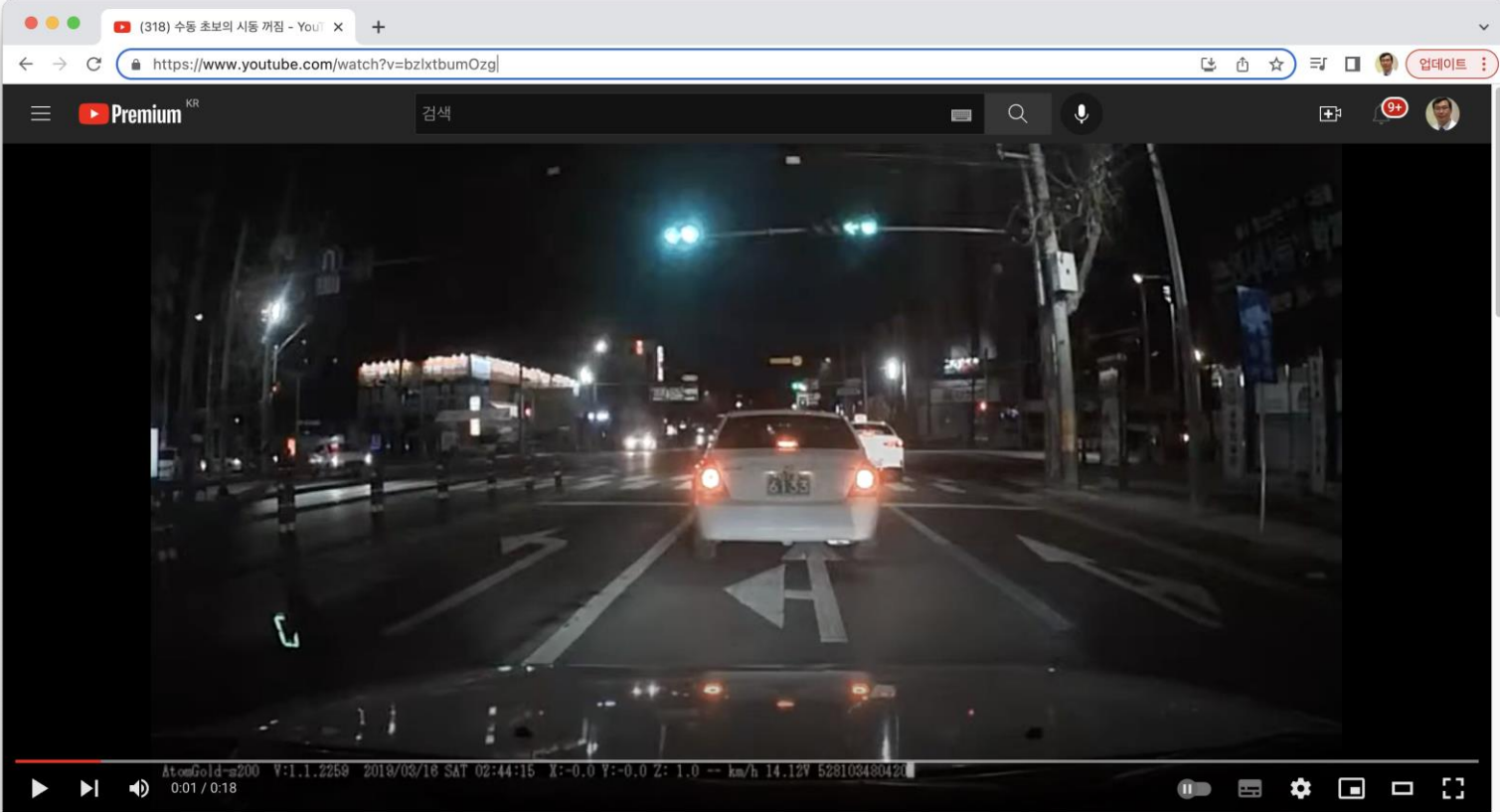
```
Backtrace:
```

```
1. base::library(MoonBook)
```

```
> |
```

<input type="checkbox"/>	.gitignore	49 B	Mar 9, 2020, 8:23 AM
<input type="checkbox"/>	.Rbuildignore	28 B	Aug 23, 2018, 2:23 PM
<input type="checkbox"/>	.Rhistory	26.7 KB	Sep 19, 2022, 12:09 PM
<input type="checkbox"/>	data		
<input type="checkbox"/>	DESCRIPTION	1.2 KB	Apr 26, 2021, 7:17 AM
<input type="checkbox"/>	inst		
<input type="checkbox"/>	man		
<input type="checkbox"/>	NAMESPACE	12.6 KB	Jul 9, 2021, 1:47 PM
<input type="checkbox"/>	NEWS.md	3.4 KB	Apr 19, 2021, 8:06 AM
<input type="checkbox"/>	R		
<input type="checkbox"/>	vignettes		
<input type="checkbox"/>	webrSub.Rproj	356 B	Sep 2, 2022, 10:09 AM

수동초보의 시동꺼짐



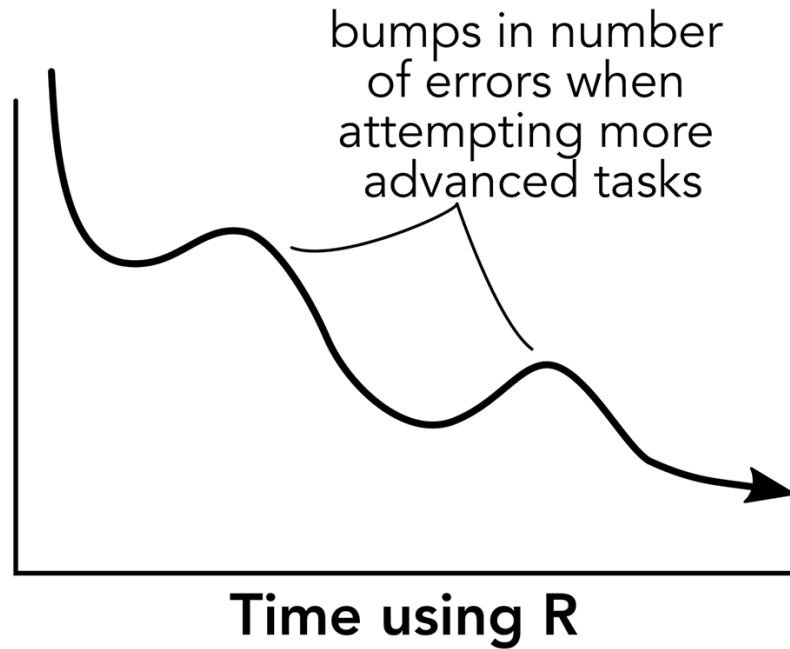
The video player shows a dashcam recording of a white car at a traffic light at night. The car's engine stalls, and the driver is seen looking back. The video title is "수동 초보의 시동 꺼짐" (Manual Beginner's Engine Stalling). The description reads: "조희수 420회 2019. 4. 16. 수동 클러치가 미흡해 시동을 두번이나 꺼트렸습니다 뒷차가 크랙신 안누르고 기다려줘서 더 미안했습니다" (Jo Heesu 420th episode 2019. 4. 16. Manual clutch was insufficient, so the engine stalled twice. The car behind cracked the horn and didn't push, so I'm even more sorry).

수동 초보의 시동 꺼짐

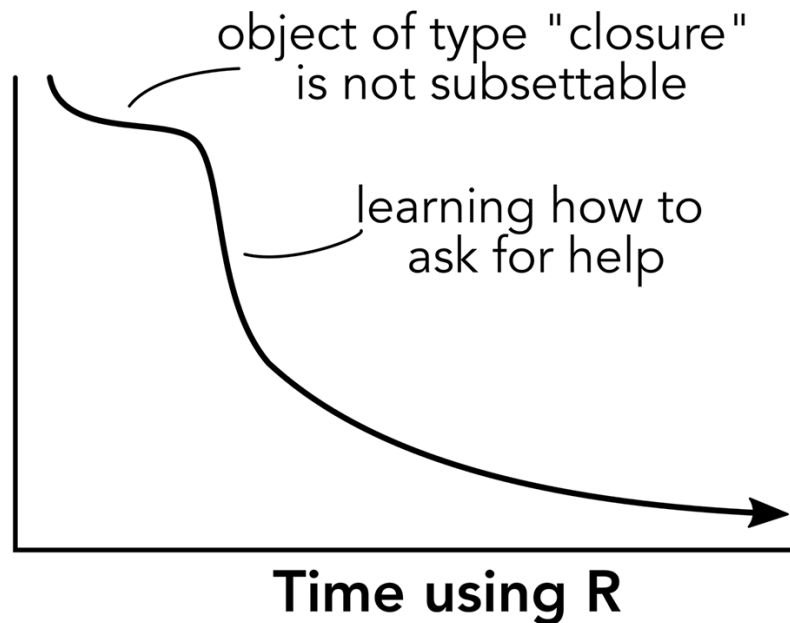
조희수 420회 2019. 4. 16. 수동 클러치가 미흡해 시동을 두번이나 꺼트렸습니다 뒷차가 크랙신 안누르고 기다려줘서 더 미안했습니다

모두 관련 콘텐츠 실시간 최근에 업로드된 동

**Number of
error messages
when writing
new code**



**Time feeling
overwhelmed
by each new
error message**



웹에서 하는 R 통계 web-r.org

The screenshot shows the homepage of web-r.org. At the top, there is a navigation bar with 'Home', '책 게시판', '게시판', '무료서버접속', '회원가입안내', and '정회원메뉴'. Below this is a header section with the title '웹에서 하는 R 통계' and a search bar. The main content area includes a welcome message, a paragraph about the website's purpose, and a list of recent posts. The '최근 로그인' section shows the user 'cardiomoon' is logged in. The '현재 접속중' section shows 'cardiomoon' is currently online. The '접속통계' section displays statistics for today, yesterday, and total visits. The '댓글작성' section lists recent comments. The '이번주 조희수' section lists recent posts by the user '조희수'. The '글작성' section lists recent posts. Below these are sections for '공지사항' (Notice), '유튜브강의게시판' (YouTube Lecture Board), and '가입인사/방명록' (Newcomer Introduction/Visitor Log).

Home | 책 게시판 | 게시판 | 무료서버접속 | 회원가입안내 | 정회원메뉴

웹에서 하는 R 통계

cardiomoon | 관리 | 로그인 | 검색 | 검색

Home | 책 게시판 | 게시판 | 무료서버접속 | 회원가입안내 | 정회원메뉴

웹에서 하는 R 통계 새로운 프로그램을 공개합니다.

"웹에서 하는 R 통계"는 통계에는 관심이 있으나 R을 어려워하는 여러 연구자들을 위한 프로젝트입니다. R설치없이 클릭만으로 웹에 있는 서버를 이용하여 통계분석을 하고 보다 R을 쉽게 사용하기 위한 패키지 개발 및 Shiny app 공동개발을 목표로 하고 있습니다. R을 사랑하는 여러분의 많은 참여를 기대합니다.

정회원께서는 로그인 하신 후 정회원용 서버 접속을 선택하여 주십시오. 처음 오신 분들은 로그인 하신후 무료서버접속 메뉴에서 웹R의 일부 기능을 체험하실 수 있습니다. 웹R 프로그램에 문제가 있는 경우 자유게시판을 통해 글 남겨주시기 바랍니다.

정회원/특별회원/단체회원으로 가입하실 분들은 회원가입안내를 읽어보시기 바랍니다.

새로 가입하신 분들은 가입인사를 남겨주세요.

cardiomoon 로그인

최근 로그인: 2022-09-15

- 회원정보 보기
- 스크린 보기
- 자강할 보기
- 작성 글 보기
- 친구 보기
- 쪽지함 보기
- 관리

현재 접속중	접속통계	댓글작성	이번주 조희수	글작성
cardiomoon	오늘: 120 어제: 809 전체: 1,473,230	1. cardiomoon ¹⁵ 2. 지미 ¹ 3. yessul ¹ 4. 빙글빙글 ¹	1. 빙글빙글 ⁸ 2. 지미 ⁷ 3. 큐레아 ⁴ 4. 햄구리 ³	1. 빙글빙글 ² 2. 지미 ¹ 3. 햄구리 ¹ 4. 큐레아 ¹
		2022-09-10 ~ 2022-09-16	2022-09-10 ~ 2022-09-16	2022-09-10 ~ 2022-09-16

공지사항

2022년 6월 17일 대한소아소화기영양학회 제 17차 워크숍 1
"의료인을 위한 R 생존분석" 책 발간 기념 이벤트 15
대한 전신알레르기 학회 서울지회 심담회 통계강좌 서버 안내 1
웹R 월간 방문수 36000 명 돌파
interpretCI 패키지가 CRAN에 등록되었습니다.
autoReg 패키지 공개
성빈센트병원 통계강좌 서버 안내 1
의정부성모병원 통계강좌 서버 안내 1
웹R 방문자수 100만명 돌파
대한중양외과학회 웨비나 서버안내 1

유튜브강의게시판

cardiomoon	PubMed 검색해서 WordCloud 만들기 5	2021-03-26 23:56	cardiomoon
cardiomoon	비교통계(1)	2021-03-12 22:30	cardiomoon
cardiomoon	기술통계, 통계 결과를 ppt나 word 형식으...	2021-03-06 12:43	cardiomoon
cardiomoon	샘분수 계산	2021-03-02 08:33	cardiomoon
cardiomoon	제1중오류와 제2중오류	2021-03-02 08:33	cardiomoon
cardiomoon	P값이란 무엇인가?	2021-03-02 08:31	cardiomoon
cardiomoon	웹에서 하는 R 통계 : 표만들기	2021-03-02 08:31	cardiomoon
cardiomoon	Propensity Score Matching 업 사용방법...	2021-03-02 08:30	cardiomoon

가입인사/방명록

등업게시판

웹에서 하는 R통계

The screenshot shows the Web-R.org website in a Chrome browser. The page title is "웹에서 하는 R 통계분석 4.2". The main content area contains a paragraph explaining that users can perform R statistical analysis online without installing R on their computer. It lists supported file formats for upload: csv, xls/xlsx, dbf, sav, dta, and sas7bdat. Below the text is a "Select Language" section with radio buttons for "English" and "한국어(Korean)", where "한국어(Korean)" is selected. A blue navigation bar lists various analysis tools: DataSelect, dataWrangling, table, Descriptive, ExploPlots, Compare, Correlation, Regression, Logistic, Poisson, Survival, Classification, Propensity Score, PPTxList, and Citation. At the bottom, there are two main sections: "파일 업로드" (File Upload) with a "Browse..." button and "데이터선택" (Data Selection) with a "데이터 전처리하기" (Data Preprocessing) section. The preprocessing section includes a checkbox for "전처리하기" (Preprocessing) and a "데이터 전처리하기" (Data Preprocessing) section with instructions on how to use R commands for data preprocessing.

Chrome 파일 수정 보기 방문 기록 북마크 프로필 탭 창 도움말 9월 20일 (화) 오후 12:19

Web-R.org cardiomoon.shinyapps.io/webr/ 일시중지됨 업데이트

웹에서 하는 R 통계분석 4.2

자신의 컴퓨터에 R을 설치할 필요 없이 R을 이용한 통계분석을 할 수 있습니다. 그룹변수와 행변수를 선택하여 쉽게 표를 만들 수 있으며 그래프를 통한 자료 탐색과 여러가지 통계분석이 가능합니다. 자신의 데이터를 xls 또는 csv형식으로 업로드하여 분석을 할 수 있을 뿐 아니라 그 결과를 pdf, docx, powerpoint 파일로 다운로드할 수 있습니다. 또한 Plot을 원하는 크기로 저장할 수 있습니다. 표가 보일 때까지 잠시만 기다려주세요.

Select Language
 English 한국어(Korean)

Web-R.org

DataSelect dataWrangling table Descriptive ExploPlots Compare Correlation Regression Logistic Poisson Survival Classification
Propensity Score PPTxList Citation

샘플 데이터를 선택 하시거나 자료를 업로드 하세요. 현재 지원하는 파일 형식은 csv, xls(x Microsoft Excel), dbf (dbase 3+), sav (SPSS), dta (STATA), sas7bdat (SAS) 파일 등이나 에러가 있을 경우 csv 형식으로 바꾸어 업로드할 것을 권장합니다.

파일 업로드
Browse... No file selected

데이터선택

전처리하기 전처리 초기화

데이터 전처리하기

여기에 R명령어를 입력하면 데이터를 전처리할 수 있습니다. R 명령어를 입력/수정하기 전에 체크박스를 해제하고 입력후 다시 체크박스를 선택하시기 바랍니다.

요약

- 통계분석(결과, 그림)을 하는 것은 우리의 목적지이다.
- SPSS, SAS를 대중교통에 비유한다면 R은 자가운전과 같다.
- R을 능숙하게 다루기 위해서는 steep learning curve를 지나야 한다.
- R에 익숙해지기 전까지 web-R.org에서 도움을 받을 수 있다.

감사합니다.