

R packing system

지금까지 R이 성장하는데 가장 큰 기여를 한 시스템 버전이 증가함에 따라 기하급수적으로 패키지 수가 증가함

R의 특징

- 해석언어 : 기계어로 변환된 뒤 실행되는 경우를 컴파일언어(compiled language)라 하고, 변환없이 바로 실행되는 경우를 해석언어(또는 인터프리트 언어(interpreted language)라 한다. C나 Java는 컴파일 언어이고 R은 인터프리트 언어이다. 컴파일 언어는 컴파일러가 코드를 프로그램으로 만든 뒤에 사용할 수 있는 반면 R은 코드를 직접 해석해서 실행한다.
- 중앙관리 : Comprehensive R Archive Network(CRAN)에서 설치파일과 팩키지를 관리한다. CRAN은 세계 여러 곳에 미러링 서버를 두고 있다.
- 클라이언트/서버 환경 지원 : 클라이언트/서버환경으로 설정해 두면 서버의 자 원을 원격으로 활용할 수 있다. 이 경우 웹 브라우저가 편집기가 된다. 서버에 작 업내용이 저장되므로 웹 브라우저를 닫아도 작업기록(history)과 환경을 저장할 수 있다

오픈소스 기반의 통계 Objected Oriented 언어

- 2000년 후반부터 Google, facebook, Amazon 등 초대형 데이터 분석에 사용
- 뉴질랜드 오클랜드(Auckland) 대학의 로스 이하카(Ross Ihaka)와 로버트 젠틀맨(Robert Gentleman) 교수의 주도하에 만들어짐
- 이들 이름의 맨 앞 글자인 이니셜을 따서 R로 명명
- 1976-1976년 Bell Laboratories에서 개발된 S 언어(statistical programming language)와 유사한 statistical programming language
- 즉, Becker 등에 의해 개발된 S 언어의 계보를 잇고 있음 [S언어를 이용한 Insightful사의 S+는 R과는 달리 S언어를 이용한 상업용 소프트웨어]
- S 언어를 기반으로 무료로 제공
- GNU S라는 별칭 : GNU's Not UNIX의 약자로 유닉스를 기반으로 누구나 자유롭게 소프트웨어를 교환하는 것을 주창하는 프로젝트 그룹 및 사상

R 개요

- Data analysis software
- A programming language
 - 통계학자들이 디자인하고 통계학자들을 위한 개발 플랫폼
- An environment
 - 데이터와 관련된 입출력, 핸들링, 관리, 분석, 그래픽 등 최신의 알고리즘 및 라이브러리 제공
- An open-source software project
 - Free, open, and active – A community
 - 수 천명의 contributors, 2백만이 넘는 사용자
 - 각 업무도메인과 관련된 리소스와 도움말 제공

-> R은 help를 쓰면 너무 잘 되어 있다.

Open source analytics for the enterprise

- R의 역량
 - Sophisticated statistical analysis
 - Predictive analytics
 - Data visualization
- R의 적용
 - Real-time trading
 - Finance
 - Risk assessment
 - Forecasting
 - Bio-technology
 - Drug development
 - Social network 등등