



R 언어 소개 및 설치 방법

환경생태데이터사이언스 실습 2019년 9월 6일

1. 오늘 실습 주제
2. R 언어 소개
3. R Studio
4. 교재

오늘 실습 주제

- GNU-R 설치
- R의 대표적인 통합 개발 환경 (Integrated Development Environment) 인 R studio 설치
- R studio 체험해보기

R 언어 소개



- The R Project for Statistical Computing
- 1993년 뉴질랜드 오클랜드 대학
- Robert Gentleman & Ross Ihaka
- 상업용 통계 언어인 S 언어 대체 목적
- 아이디어 시험 플랫폼 → 교육 도구 → 연구 및 개발 목적

그럼 왜 R이라는 이름이 붙었을까요?

- 통계 연산과 도표 작성을 주된 목표로 하는 프로그래밍 언어

- 통계 연산과 도표 작성을 주된 목표로 하는 프로그래밍 언어
- 다양한 도구들을 제공하는 소프트웨어 환경

- 통계 전문 언어인 S 언어에서 파생
 - 통계 처리 및 도표 작성에 강점
- 교육 목적으로 이용
 - 사용하기 쉬운 언어

- Free and Open-source software (FOSS)
 - 무료로 자유롭게 사용 가능
- 연구 기관과 산업 분야에 종사하는 많은 사용자
 - 문제가 생겼을 때 도움 받기가 쉬움.
- 다수의 개발자들이 자발적으로 패키지 제작 및 관리.
 - 새로운 통계 기법 및 자료 처리 기법들이 빠르게 적용됨.

R의 특징

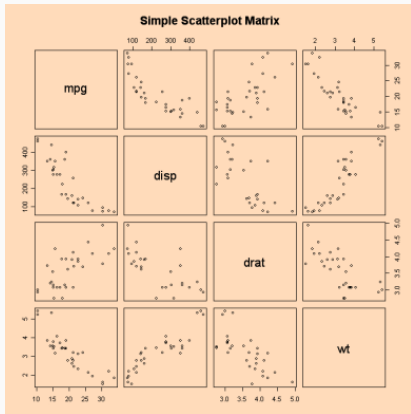
Worldwide, Sept 2019 compared to a year ago:

Rank	Change	Language	Share	Trend
1		Python	29.21 %	+4.6 %
2		Java	19.9 %	-2.2 %
3		Javascript	8.39 %	+0.0 %
4		C#	7.23 %	-0.6 %
5		PHP	6.69 %	-1.0 %
6		C/C++	5.8 %	-0.4 %
7		R	3.91 %	-0.2 %
8		Objective-C	2.63 %	-0.7 %
9		Swift	2.46 %	-0.3 %
10		Matlab	1.82 %	-0.2 %
11	↑	TypeScript	1.77 %	+0.2 %
12	↑↑↑↑↑	Kotlin	1.55 %	+0.6 %
13		VBA	1.44 %	-0.0 %
14	↓↓↓	Ruby	1.4 %	-0.2 %
15	↑↑	Go	1.2 %	+0.3 %
16	↓↓	Scala	1.14 %	-0.1 %

R의 특징

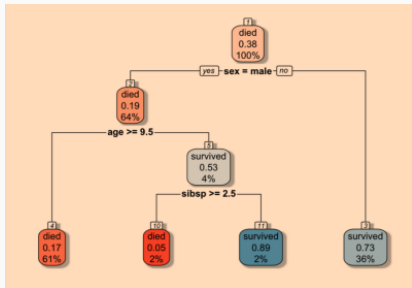
Aug 2019	Aug 2018	Change	Programming Language	Ratings	Change
1	1		Java	16.028%	-0.85%
2	2		C	15.154%	+0.19%
3	4	▲	Python	10.020%	+3.03%
4	3	▼	C++	6.057%	-1.41%
5	6	▲	C#	3.842%	+0.30%
6	5	▼	Visual Basic .NET	3.695%	-1.07%
7	8	▲	JavaScript	2.258%	-0.15%
8	7	▼	PHP	2.075%	-0.85%
9	14	▲▲	Objective-C	1.690%	+0.33%
10	9	▼	SQL	1.625%	-0.69%
11	15	▲▲	Ruby	1.316%	+0.13%
12	13	▲	MATLAB	1.274%	-0.09%
13	44	▲▲	Groovy	1.225%	+1.04%
14	12	▼	Delphi/Object Pascal	1.194%	-0.18%
15	10	▼▼	Assembly language	1.114%	-0.30%
16	19	▲	Visual Basic	1.025%	+0.10%
17	17		Go	0.973%	-0.02%
18	11	▼▼	Swift	0.890%	-0.49%
19	16	▼	Perl	0.860%	-0.31%
20	18	▼	R	0.822%	-0.14%

R의 용도 - 통계 분석



출처: Quick R

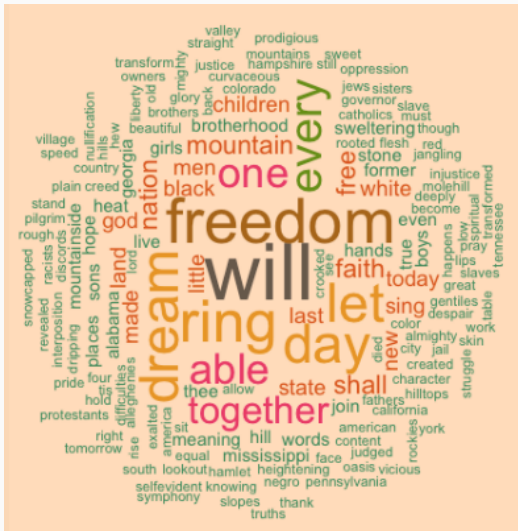
- 기초 통계
 - 상관계수 분석
 - 선형 분석
- 고급 통계
 - 베이지안 분석
 - 비선형 분석



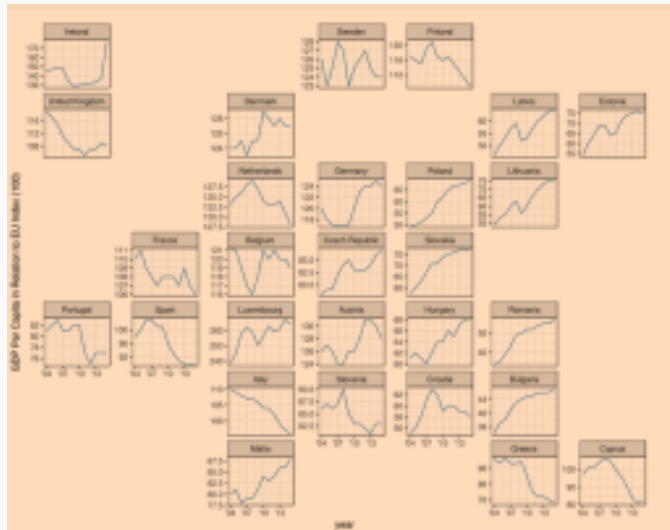
출처: Kei Saito (2015) Visualizing a decision tree using R packages in Exploratory

- 다량의 자료 이용
- 특정 변수 예측
- 특징 파악
 - Random forest (RF)
 - Support vector machine (SVM)
 - Deep learning (DL)

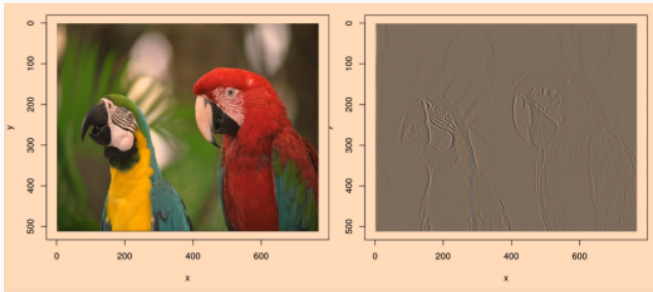
R의 용도 - 기타



R의 용도 - 기타



R의 용도 - 기타



출처: imager: an R package for image processing

R Studio

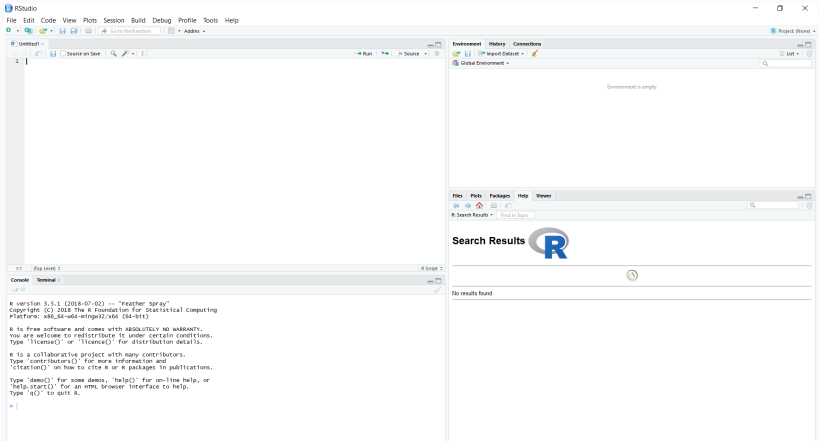




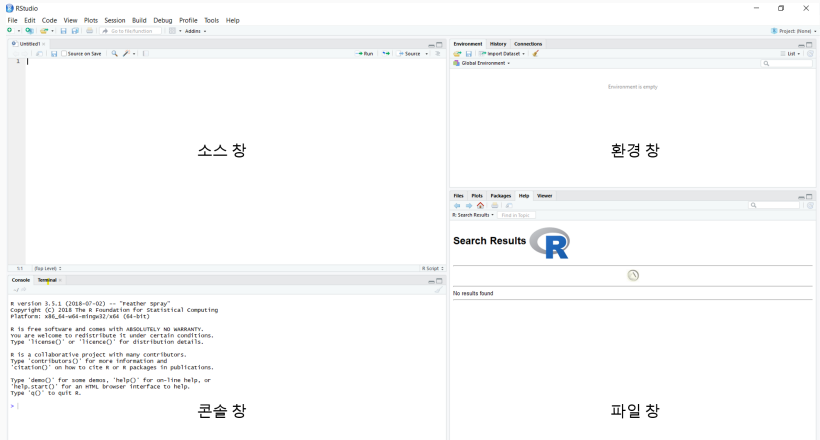
- R studio 제작
- R 언어를 위한 가장 대중적인 통합 개발 환경
- 제한적인 공개 프로그램
- R project의 발전을 위해 다양한 기여

- 가장 쉽고 편리한 R 전용 통합 개발 환경
- 코드 재사용 용이
- 프로젝트 관리 도구와 문서 작성 도구 포함
- R Studio server → 서버/클라우드 시스템에서도 편하게 작업 가능

R Studio의 기본 창 구성



R Studio의 기본 창 구성



- 코드 작성
- 콘솔 창으로 코드를 보냄 [Ctrl + enter]
- 코드 재사용을 위해 필수적으로 이용

- 코드 실행
- 간단한 테스트 코드는 콘솔 창에 직접 입력 가능

- 코드 실행 결과 생성된 자료 정보
- 변수, 벡터, 행렬, 테이블

- Files → 현재 작업 중인 폴더의 디렉토리 정보
- Plots → 그래프 표시
- Packages → 설치된 패키지 목록
- Help → help() 함수의 실행 결과 (도움말)
- Viewer → 분석 결과를 HTML 등 웹 문서로 출력한 모습을 보여줌

교재

- “R을 이용한 데이터 처리 & 분석 실무”

- 무료로 도서 공개 (pdf)

: http://r4pda.co.kr/pdf/r4pda_2014_03_02.pdf

- 무료로 도서 공개 (html)

: <https://thebook.io/006723/>