



미래창조과학부

**NIA**

한국정보화진흥원  
NATIONAL INFORMATION SOCIETY AGENCY

빅데이터 분석전문가 양성을 위한

# 빅데이터 분석 실습 교육 - R 기초

**2e** 투이컨설팅



미래창조과학부

**NIA** 한국정보화진흥원  
NATIONAL INFORMATION SOCIETY AGENCY

- I. R 소개
- II. R에서 자료 타입에 따른 Data manipulation
- III. 데이터 입출력
- IV. 간단한 도표를 이용한 자료 요약
- V. 프로그래밍

# Chapter I. R 소개

BIG  
DATA



# R 소개

---

- 본 장에서는 최근 데이터 과학에서 공용어로 자리잡아가는 R에 대해서 가볍게 소개한다.
- Issue

R실행시 처음으로 만나게 되는 Console 과 Script 창에 대한 설명

작업 디렉토리(Working directory)와  
패키지(Packages)에 대한 이해

산술연산을 하기위한 기본적인 연산자와 변수생성법, 주석표기법

# Chapter II. R에서 자료 타입에 따른 DATA MANIPULATION



# R에서 자료 타입에 따른 Data manipulation

---

- 본 장에서는 데이터 타입에 따른 자료 생성 및 조작법의 기초를 익힌다.
- Issue

데이터 타입은 여러 가지가 있음

vector, 배열(array), 행렬(matrix), list, data.frame,  
factor, 문자열(string), 날짜(Date) 이상 8개를 알고 넘어가야 함

데이터 타입은 서로 변환이 가능하며, 이점을 이용하여 입맛에 맞게  
데이터 핸들링을 할 수 있다는 점을 파악

# Chapter III. 데이터 입출력



# 데이터 입출력

---

- 본 장에서는 데이터를 입력하는 방법, 외부에 다른 포맷으로 준비된 자료를 불러오는 방법, 반대로 내보내는 방법에 대해 학습한다.
- Issue

데이터 포맷에 따라 그에 맞는 함수들을 사용하여야 함

Foreign package 를 이용하여 SPSS, SAS 등에서 편집된 데이터셋도 읽어들이 수 있음



# Chapter IV . 간단한 도표를 이용한 자료요약



# 간단한 도표를 이용한 자료 요약

---

- 자료의 성격에 따라 적절한 도표 작성법을 사용하여야 의미 있는 정보가 담긴 도표를 작성할 수 있다.

- Issue

이전에 질적자료와 양적자료의 성격을 파악하고  
자료 형태에 따라 의미있는 정보가 담길 수 있도록 노력하여야 함

질적자료는 Bar chart, Pie chart 가 대표적

양적자료는 Histogram, Boxplot, Scatterplot 등이 대표적

# Chapter V .프로그래밍

BIG  
DATA



# 프로그래밍

---

- 반복적인 작업을 단순화 하거나 코드를 보다 읽기 좋도록 하기 위하여 R 프로그래밍에서 사용하는 제어문 들을 익힌다.
- Issue

조건문과 반복문의 제어문들에 대한 개념을 파악

제어문들과 사용자 함수를 결합하여  
좀 더 의미있게 반복작업을 줄이기 위한 응용

# Q&A



미래창조과학부

**NIA**

한국정보화진흥원  
NATIONAL INFORMATION SOCIETY AGENCY

**2e** 투이컨설팅