

R을 이용한 통계 기초와 데이터 분석

한성대학교
강사: 남현진
jinanam0116@gmail.com

강의 소개

이 강의는 통계 및 데이터 분석에 관심이 있는 사람들이 빠르게 실무에 이용되는 개념들을 이해하고 R을 이용하여 실습해 볼 수 있는 통계 입문자용 강의입니다. 데이터 분석에는 굉장히 다양한 주제와 기술들이 있습니다. 이 시대에 살아가는 모두는 4차 산업혁명과 빅데이터라는 키워드에는 익숙하지만 실제로 그 기술이 무엇인지, 산업에서 필요시 되는 능력이 무엇인지 이해하고 알아가려 하면 시작이 어디인지 막막합니다. 데이터 분석은 거창한 것들 뿐만 아니라 누구나 배우면 바로 사용할 수 있는 실용적인 기술들로부터 시작합니다. 이 강의에서는 데이터 분석이 무엇인지 또 어떤 분야가 존재하는지 이해하고 현업에서 사용되는 간단한 통계 지식을 배우려 합니다. 통계 분석, 시각화, 기계 학습에 특화되어 있는 통계 지식과 R언어의 기초를 배우고 짧은 시간 안에 실전 문제를 적용하여 데이터 분석이 무엇인지 알아보려 합니다.

강의 자료 및 준비사항

강의 자료는 PPT로 제공됩니다. 실습은 프로그래밍 R을 사용하며 R과 RStudio를 설치 할 수 있는 컴퓨터가 필요합니다.

참고 문헌

- Norman Matloff (2011), The Art of R Programming: A Tour of Statistical Software Design, No Starch Press.
- Grolemund, Garrett and Wickham, Hadley (2016), R for Data Science, O'Reilly
- 서민구 (2014), R을 이용한 데이터 처리 분석 실무, 길벗

강의 내용

주차	이론 및 실습
1 주차	<ul style="list-style-type: none">- 기초 통계의 개념 (평균, 중앙값, 분산, 행렬, 확률 분포)- R설치 및 인터페이스 설명- 실습: R로 평균과 분산 계산하기
2 주차	<ul style="list-style-type: none">- 데이터 다루기- 기본 함수(For, Ifelse, while, function)- 실습: 최소값/최대값을 출력하는 알고리즘 작성하기
3 주차	<ul style="list-style-type: none">- 데이터 불러오기- Ggplot2 패키지를 이용한 데이터 시각화- 실습: Covid-19 데이터를 이용한 데이터 시각화
4 주차	<ul style="list-style-type: none">- 통계적 가설 (귀무가설, 대립가설)- P-value, 유의수준, 통계의 오류 (제 1종 오류, 제 2종 오류)- 추정 및 신뢰구간
5 주차	<ul style="list-style-type: none">- T-test- 실습1 : 남자와 여자 간 소득의 차이 비교- 실습2 : 교육을 받기 전과 후의 반 학생들의 성적 변화
6 주차	<ul style="list-style-type: none">- 선형회귀분석- 실습: 자동차 주행 속도를 이용해 제동 거리를 예측하기
7 주차	<ul style="list-style-type: none">- 로지스틱 회귀분석- 실습: 의학 데이터를 사용한 당뇨병 진단 모형 만들기