



2021 학년도 2학기 강의계획안

교과목명 Course Title	빅데이터를 이용한 통계 그래픽스	학수번호 Course No.	38191(02)
개설전공 Department/Major	통계학과	학점 Credit	3
수업시간/강의실 Class Time/ Classroom	수 15:30 ~ 16:45, 금 12:30 ~ 13:45 / 비대면수업		
담당교원 Instructor	성명: 이 은 경 Name	소속: 통계학과 Department	
	E-mail: lee.eunk@ewha.ac.kr	연락처: 02-3277-6857 Telephone	
면담시간/장소 Office Hours/ Office Location			

I. 교과목 정보 Course Overview

1. 교과목 개요 Course Description

본 강의에서는 빅 데이터를 이해하고 정리하고, 이를 시각적으로 분석하여 새로운 가치 창출에 사용하기 위한 전 과정을 학습하는 것을 목표로 한다. 데이터는 기하급수적으로 늘어가고 있지만 분석과 새로운 가치를 만들어 낼만한 기술적, 전략적 분석 방법들은 상대적으로 미흡한 것이 현실이다. 빅 데이터 분석을 위해서는 창의력과 논리적 판단이 필요하며 기존의 데이터 분석 방법론으로는 의미있는 결론을 도출하기 어렵다. 본 강의에서는 통계 자료 탐색의 기본이 되는 자료정리 기법 뿐 아니라 통계 그래픽스 기법들을 창조적으로 활용하여 빅 데이터로부터 의미 있는 결과를 도출하는 방법을 학습하게 된다. 이는 기술이 아니라 창조와 논리적 사고의 영역인 빅 데이터 분석을 위한 기초 교육에 꼭 필요한 부분으로 통계 그래픽스를 이용한 빅 데이터 분석을 통하여 빅 데이터를 적극적으로 활용할 수 있게 된다.

2. 선수학습사항 Prerequisites

통계학 II 또는 확률및통계학, 전산통계및실습, 컴퓨터 프로그래밍(36856)



3. 강의방식 Course Format

강의 Lecture	발표/토론 Discussion/Presentation	실험/실습 Experiment/Practicum	현장실습 Field Study	기타 Other
50 %	20 %	30 %		%

(위 항목은 실제 강의방식에 맞추어 변경 가능합니다.)

강의 진행 방식 설명 (explanation of course format):

본 강의는 수강인원 50명 이상의 이론/실습 교과목으로 학교의 방침에 따라 전 기간 비대면 수업(온라인 수업)으로 진행됨.



1. 강의록과 강의 동영상은 **매주 수요일, 금요일 6:00AM**에 사이버캠퍼스에서 제공됨. 학생들은 다음 수업시간 전까지 동영상을 들어야 함.
2. **매주 금요일 12:30 PM ~ 13:15 PM**에는 ZOOM을 이용한 온라인 질의응답이 진행됨. (참여율이 성적에 반영됨)
3. 중간, 기말고사는 대면시험을 기본으로 하나 상황에 따라 온라인 시험으로 대체될 수 있음.
4. 프로젝트 발표는 대면발표를 기본으로 하나 상황에 따라 조별로 VOD를 찍거나 ZOOM에서 발표를 하는 형태로 대체될 수 있음.
5. 온라인 수업이므로 **공휴일 없이 진행됨**.

4. 교과목표 Course Objectives

본 강의에서는 빅 데이터의 기본 개념과 자료의 형태에 대하여 학습하고 빅 데이터를 위한 데이터 베이스에 대하여 살펴본다. 또한 빅 데이터의 데이터베이스로부터 자료를 추출하여 유용한 정보를 분석할 수 있는 형태로 정리하고 이를 자료의 특성이 잘 나타나는 형태의 그림으로 표현할 수 있는 통계 그래픽스 방법을 학습한다. 이는 시각적으로 자료를 탐색하고 분석하는 다양한 방법으로 이론 뿐만 아니라 실습을 통하여 실제 빅 데이터로부터 유용한 정보를 시각적으로 정리할 수 있는 능력을 기른다. 이 모든 과정을 위하여 통계 언어인 R을 이용한다. R의 기초문법을 습득하고 이를 기반으로 하여 대용량의 자료를 정리하는 기법을 학습한다. 또한 R을 기본으로 하여 자료를 다양한 형태로 탐색할 수 있는 그래픽 기법들을 학습하고 이를 대용량 자료에 적용시켜본다. 또한 대용량 자료 그래픽에서 발생하는 문제점들을 인지하고 이에 대한 해결 방안을 학습한다. 자료 분석 결과를 원 자료와 접목하여 그래픽으로 표현하여 자료의 특성을 좀 더 잘 표현할 수 있는 방법을 학습한다. 실제 대용량 자료를 이용하여 자료의 정리부터 자료를 탐색하는 과정, 그리고 이를 보고서로 작성하고 이를 발표해 봄으로써 자료 정리, 탐색의 내용을 효율적으로 전달할 수 있는 능력을 기른다.



5. 학습평가방식 Evaluation System

■ 상대평가(Relative evaluation) □ 절대평가(Absolute evaluation) □ 기타(Others): _____

- 평가방식 설명 (explanation of evaluation system):

* 아래의 기준에 따른 상대평가임

A등급 (A+, A0, A-)	35% 이내
A등급과 B등급(B+, B0, B-)의 합	70% 이내
C등급 이하(C+, C0, C-, D+, D0, D-, F)	30% 이상

* 재수강 학생들의 학점은 B-부터 시작됨.

* 온라인 시험으로 변경되는 경우 절대평가로 바뀔 수 있음

중간고사 Midterm Exam	기말고사 Final Exam	퀴즈 Quizzes	발표 Presentation	프로젝트 Projects	과제물 Assignments	참여도 Participation	기타 Other
25 %	35 %	5 %	%	10 %	20 %	5 %	%

(위 항목은 실제 학습평가방식에 맞추어 변경 가능합니다.)

* 매주 금요일 takehome homework이 주어짐

* 중간고사 : 10월 20일 (수) 19:00 ~ 21:00 (대면시험 예정)

* 기말고사 : 12월 1일 (수) 19:00 ~ 21:00 (대면시험 예정)

* 대면시험을 기본으로 하나 COVID-19의 상황에 따라 온라인 시험으로 대체될 수 있음.

* 대면/온라인 시험 모두 시험시간은 변경되지 않으며 조정 불가능하니 꼭 확인 후 수강신청을 해야 함

* 출석은 다음 시간 전까지의 동영상 진도현황과 금요일의 질의응답 참여도로 이루어짐.

* 전체수업의 1/3 이상을 결석할 경우 학칙에 따라 낙제처리 됨.

* 과제물을 위해 학생들 간에 논의하는 정도는 허용되나 학생들 각각은 각자의 고유 프로그램 코드를 제출하여야 함.

* 프로젝트는 팀별 프로젝트로 기말고사 후에 시행될 예정임.

* 팀 단위로 진행되는 프로젝트의 팀원 구성원은 랜덤으로 정해지며 중간고사 이후 발표됨.



Ⅱ. 교재 및 참고문헌 Course Materials and Additional Readings

1. 주교재 Required Materials

교재는 없으며 모든 강의 자료는 cyber campus에서 제공됨

2. 부교재 Supplementary Materials

R for data science - Hadley Wickham

ggplot2 - H. Wickham - Springer

3. 참고문헌 Optional Additional Readings

비주얼라이즈 디스 - 네이션 야우 지음 / 송용근 옮김 - 에이콘

R을 이용한 빅데이터분석 - 임동훈 - 자유아카데미

빅데이터 분석도구 R프로그래밍 - 노만 매트로프 지음 - 에이콘

The R Book - M. J. Crawley - Wiley

Dynamic Documents with R and knitr - Yihui Xie - CRC Press

Reproducible Research with R and Rstudio - C. Gandrud - CRC Press

Advanced R - H. Wickham - CRC Press

Data manipulation with R - P. Spector - Springer

Ⅲ. 수업운영규정 Course Policies

*

- * 실험실 진행 강의의 경우 본교에서 진행되는 '실험실안전교육'을 필수로 이수하여야 함.
- * 3번의 지각은 1번의 결석으로 처리됨.
- * 10번 이상의 결석시 F처리됨.



IV. 주차별 강의계획 Course Schedule (최소 15주차 이상 강의)

주차	날짜		
1주차	9월1일 (수)/ 9월3일 (금)/	주요강의내용 및 구성 Topics & Class Format	Introduction to big data
		주요강의내용 및 구성 Topics & Class Format	How to analyze & visualize big data in data base
2주차	9월8일 (수)/ 9월10일 (금)/	주요강의내용 및 구성 Topics & Class Format	Basic grammar in R (1)
		주요강의내용 및 구성 Topics & Class Format	Basic grammar in R (2)
3주차	9월15일 (수)/ 9월17일 (금)/	주요강의내용 및 구성 Topics & Class Format	ggplot2 - intro
		주요강의내용 및 구성 Topics & Class Format	ggplot2 - toolbox
4주차	9월22일 (수)/ 9월24일 (금)/	주요강의내용 및 구성 Topics & Class Format	ggplot2 - grammar and layer
		주요강의내용 및 구성 Topics & Class Format	ggplot2 - scales, axes and legends
5주차	9월29일 (수)/ 10월1일 (금)/	주요강의내용 및 구성 Topics & Class Format	ggplot2 - positioning
		주요강의내용 및 구성 Topics & Class Format	ggplot2 - theme
6주차	10월6일 (수)/ 10월8일 (금)/	주요강의내용 및 구성 Topics & Class Format	ggplot2 - data transformation
		주요강의내용 및 구성 Topics & Class Format	ggplot2 - modeling and programming
7주차	10월13일 (수)/ 10월15일 (금)/	주요강의내용 및 구성 Topics & Class Format	ggplot2 - case study (1)
		주요강의내용 및 구성 Topics & Class Format	ggplot2 - case study (2)
8주차	10월20일 (수)/ 10월22일 (금)/	주요강의내용 및 구성 Topics & Class Format	Review
		주요강의내용 및 구성 Topics & Class Format	How to handle big data with R - intro



주차	날짜		
9주차	10월27일 (수)/ 10월29일 (금)/	주요강의내용 및 구성 Topics & Class Format	data transformation
		주요강의내용 및 구성 Topics & Class Format	Tibbles and data import
10주차	11월3일 (수)/ 11월5일 (금)/	주요강의내용 및 구성 Topics & Class Format	Tidy data in R
		주요강의내용 및 구성 Topics & Class Format	strings/factors
11주차	11월10일 (수)/ 11월12일 (금)/	주요강의내용 및 구성 Topics & Class Format	Shiny (1)
		주요강의내용 및 구성 Topics & Class Format	Shiny (2)
12주차	11월17일 (수)/ 11월19일 (금)/	주요강의내용 및 구성 Topics & Class Format	Shiny (3)
		주요강의내용 및 구성 Topics & Class Format	Rmakrdown
13주차	11월24일 (수)/ 11월26일 (금)/	주요강의내용 및 구성 Topics & Class Format	text mining with visualization
		주요강의내용 및 구성 Topics & Class Format	sentiment analysis with visualization
14주차	12월1일 (수)/ 12월3일 (금)/	주요강의내용 및 구성 Topics & Class Format	Review
		주요강의내용 및 구성 Topics & Class Format	team project presentation (1) (대면발표)
15주차	12월8일 (수)/ 12월10일 (금)/	주요강의내용 및 구성 Topics & Class Format	team project presentation (2) (대면발표)
		주요강의내용 및 구성 Topics & Class Format	team project presentation (3) (대면발표)
시험		중간고사 (대면시험)	10월 20일 (수) 19:00 ~ 21:00
		기말고사 (대면시험)	12월 1일 (수) 19:00 ~ 21:00



V. 참고사항 Special Accommodations

- * 학칙 제57조에 의거하여 장애학생은 학기 첫 주에 교과목 담당교수와의 면담을 통해 출석, 강의, 과제 및 시험에 관한 교수학습 지원 사항을 요청할 수 있으며 요청된 사항에 대해 담당교수 또는 장애학생지원센터를 통해 지원받을 수 있습니다.

According to the University regulation #57, students with disabilities can request special accommodation related to attendance, lectures, assignments, and/or tests by contacting the course professor at the beginning of semester. Based on the nature of the students' requests, students can receive support for such accommodations from the course professor and/or from the Support Center for Students with Disabilities (SCSD).

- * 강의계획안의 내용은 추후 변경될 수 있습니다.
- * 02 분반은 3학년 이상의 고학년을 위한 수업으로 2학년 학생이 등록을 하는 경우 학점에 불이익을 받을 수 있습니다. 2학년 학생들은 꼭 01 분반으로, 3학년 이상의 고학년 학생들은 02 분반으로 신청해야 합니다.