

326.212: 전산통계 및 실습

강의계획서, 2019 가을학기

Course Home Page: https://won-j.github.io/326_212-2019fall/

강의목표

본 과목에서는 통계적 프로그래밍을 통해 자료를 수집, 가공, 분석 및 프리젠테이션하는 방법을 공부한다.

이 과목에서 배우는 내용은 크게 다음과 같다.

- Basic programming skills using the R programming language
- Proficiently wrangle, manipulate, and explore data using R
- Utilize contemporary R packages, especially the tidyverse
- Visualize, present, and communicate trends in a variety of data types
- Formulate data-driven hypotheses using exploratory data analysis and introductory model building techniques

목표 달성을 위해 다수의 프로그래밍 숙제와 프로젝트를 수행하며, 수강생은 이를 위해 수업 외 주당 10시간 정도의 노력이 필요할 것으로 예상된다.

강사 및 조교

강사:

원중호 25동 329호, 880-4121 wonj@stats.snu.ac.kr

조교:

이성동 25동 304호, 880-4122 clan1212@snu.ac.kr

최영원 25동 304호, 880-4122 muha@snu.ac.kr

강의시간 및 강의실

강의실:

28동 101호 월요일 13:00 – 14:50 (강의)

실습실:

26동 102호 수요일 13:00 – 14:50 (실습)

선수과목

통계학 (033.019), 통계학 실험 (033.020) 또는 그 상당 과목.

교재 및 참고문헌

- Hadley Wickham and Garret Golemund. R for Data Science. O'Reilly Media, 2017. ISBN: 978-1491910399. (교재).
- Norman Matloff. The Art of R Programming. No Starch Press, 2011. ISBN: 1593273843 (참고문헌).
- Hadley Wickham. Advanced R. CRC Press, 2014. ISBN: 978-1466586963. (참고문헌).

평가방법

본 과목의 평가 방법은 아래와 같다.

- 수업참여 (10%). 수강 태도 및 수업에 적극적으로 참여하는 정도. **수업일수의 1/3을 초과하여 결석하면 성적은 “F”가 됨** (담당교수가 불가피한 결석으로 인정하는 경우는 예외로 할 수 있음). 학기 종료 후 정당한 사유 없이 학점을 올려달라고 요청할 경우 수업참여도 점수를 0점으로 처리하여 불이익을 받을 수 있다. 마찬가지로 재수강을 이유로 학점 변경을 요청하는 경우도 허용하지 않는다.
- 숙제 (30%). 교재의 문제 풀이와 간단한 프로그래밍. 수시로 부과되며, 숙제는 마감일 오후 11시 59분까지 채점용 github (추후 공지)에 온라인으로 제출해야 한다. 작성한 코드는 github에 바로 제출할 수 있으며, 수반되는 문서는 Markdown으로 작성하여야 한다. 늦게 제출한 숙제는 1일당 전체 점수의 50%씩 감점한다.

- 퀴즈 (20%). 실습시간에 수시로 프로그램 작성 퀴즈를 실시한다. 제출은 숙제와 마찬가지로 채점용 github에 온라인으로 하며, 정해진 시간 안에 완료해야 한다.
- 기말 프로젝트 (40%). 실제적인 문제 해결을 위한 프로그래밍 프로젝트. 제출 방식은 숙제의 경우와 같다. 프로젝트의 자세한 개요는 추후 공지하나, 필요한 자료의 크기가 30 기가바이트를 넘어갈 것으로 전망되므로, 이에 맞도록 각자 컴퓨터의 하드디스크 공간을 준비할 것을 권장한다.

숙제 협력에 관한 규칙

수업을 듣는 동료 학생들과 같이 숙제를 하는 것은 장려한다. 하지만 답안지는 반드시 본인 스스로 작성해야 하며 도움을 받을시에는 도움을 준 학생의 이름을 숙제 제출시 명시해야한다. 그렇지 않을 경우 같은 답안을 제출한 숙제는 모두 0점 처리한다.

숙제, 프로젝트 및 시험 중 어느 곳에서라도 부정행위가 발각될 경우, 해당 수강자의 학점을 F 처리하며, 해당 수강자는 징계위원회에 회부될 수 있다. 특히, **다른 사람의 프로그램 코드를 베껴서 제출하는 행위는 심각한 부정행위임**을 명시한다.

강의계획

1주차 (9/2, 9/4) Introduction, Data Visualization I

2주차 (9/9, 9/11) Workflows, R Markdown

3주차 (9/16, 9/18) Data Visualization II, Data Transformation I

4주차 (9/23, 9/25) Data Transformation II

5주차 (9/30, 10/2) Exploratory Data Analysis

6주차 (10/7) Import and Tidy Data I

7주차 (10/14, 10/16) Import and Tidy Data II

8주차 (10/21, 10/23) Relational Data

9주차 (10/28, 10/30) Strings, Factors, Date and Times

10주차 (11/4, 11/6) Vectors

11주차 (11/11, 11/13) Functions, Pipes

12주차 (11/18, 11/20) Iteration

13주차 (11/25, 11/27) Model Basics

14주차 (12/2, 12/4) Model Building

15주차 (12/9, 12/11) 기말 프로젝트

위 계획은 수업 진도에 따라 변경될 수 있다.