

L0. Introduction

Sim, Min Kyu, Ph.D., mksim@seoultech.ac.kr



서울과학기술대학교 데이터사이언스학과

강의 정보

- 데이터 시각화 (9490012)
- 2020 Spring
- 월요일 10시-13시, 프론티어관 108호
- 빅데이터를 효과적으로 시각화하는 기법을 학습한다.
- Summary
 - ▶ 기본적인 데이터의 시각화 디자인 기법과 평가 방법
 - ▶ 다변량/텍스트/네트워크 데이터 등 다양한 형태의 데이터에 대한 시각화
 - ▶ R 언어를 통해 구현 및 검증한다.
 - ▶ 데이터사이언스학과 석사 2,3,4학기 학생을 대상으로 한다.
 - ▶ (해당학기가 아닌데 수강을 원하는 경우에는 면담요청할것)
- Goal
 - ▶ 강력한 시각화 도구인 R의 사용에 능숙해진다.
 - ▶ 의사소통의 목적에 적합한 시각화 결과물을 생산할 수 있다.
 - ▶ 다양한 종류의 데이터의 정보를 요약하여 시각화 결과물을 생산할 수 있다.
 - ▶ 논문발표/학술대회/인터뷰 등의 상황에서 훌륭한 프리젠테이션을 할 수 있다.

- 심민규

- ▶ `mksim@seoultech.ac.kr`
- ▶ 프론티어관 915호
- ▶ 개인 면담
 - 수업 전후/이메일로 약속을 잡으면 됩니다.
 - 주로 월요일과 수요일 오후에 면담을 하기 좋습니다.

- 교재

- ▶ 해들리 위컴, 개릿 그롤문드 저/김설기, 최혜민 역. 'R을 활용한 데이터 과학' (2019) (번역본)
 - 과제와 Reading Assignment등 주교재
 - Grolemond, Garrett, and Hadley Wickham. 'R for data science.' (2018) (영문원서)
- ▶ 강의노트
- ▶ `datacamp.com`의 온라인 튜토리얼

- 보조 교재

- ▶ 윈스턴 쉥 저/이제원 역. 'R Graphics Cookbook:데이터 시각화를 위한 실용 레시피' (번역본)
- ▶ Chang, Winston. R graphics cookbook: practical recipes for visualizing data. O'Reilly Media, 2018. (위의 도서의 영문원서)
- ▶ Tufte, Edward R. The visual display of quantitative information. Vol. 2. Cheshire, CT: Graphics press, 2001.

Your work and Grading

1 읽기 과제 [ra]

- ▶ 수업전에 부여된 분량을 미리 읽고 올 것

2 datacamp.com [dc]

- ▶ 부여된 module을 완료할 것
- ▶ 학점에서의 비중은 20%

3 수업 중 토론

- ▶ 읽기 과제와 사전 지식을 바탕으로 수업중 토론에 임하며,
- ▶ 타 학우의 작업물에 대한 좋은 품질의 피드백을 제공
- ▶ 학점에서의 비중은 10%

4 과제 [ps]

- ▶ 문제풀이와 시각화 결과물 구현의 형태로 주어지며,
- ▶ 2명의 학우가 팀이 되어 같이 수행함
- ▶ 팀은 계속 random하게 assign할 예정
- ▶ 학점에서의 비중은 50%

5 기말 프로젝트

- ▶ 가급적이면 본인의 연구와 연관된 주제를 선정해서 시각화 결과물을 발표
- ▶ 10분간 개인별 발표
- ▶ 학점에서의 비중은 20%

Your work and Grading

- 과제에 시간 투자가 필요하지만, 중간/기말 고사가 없음
- 수강생들의 아래의 skill set을 키우는데에 초점을 두고 수업을 진행함
 - ▶ 데이터를 시각화 skills을 이용해서 정리할 수 있다.
 - ▶ 시각화한 결과물을 이용하여 written/verbal communication의 형태로 효과적으로 타인과 공유할 수 있다.
- [dc]와 [ps]에 대해서는 n 개의 과제가 부여되었다면, $n-1$ 개를 이용해서 학점을 부여함
- 수업과 관련해서 프로페셔널하고 신사/숙녀답게 행동할 것을 부탁드립니다.
 - ▶ 팀으로 수행하는 과제에서 충실히 책임을 다하고
 - ▶ 발표나 피드백의 교환에 있어서 프로페셔널하고
 - ▶ 타인과 본인을 항상 존중하는 자세를 보여주세요.
- 데이터 시각화는 보고서와 프리젠테이션의 도구입니다. 과제물등을 professional and neat하게 준비하세요.
- Resource로 남기고자 수강생들의 work를 모은 github를 운영할 예정입니다.

Topics

- Part I. R's tidyverse
 - ▶ Module 1 (R basics)
 - ▶ Module 2 (Wrangling)
 - ▶ Module 3 (ggplot2)
- Part II. 데이터 형식과 시각화
 - ▶ Module 4 (text data)
 - ▶ Module 5 (geospatial data)
 - ▶ Module 6 (time-series data)
- Part III. 대시보드
 - ▶ Module 7 (dashboard)
 - ▶ Module 8 (shiny)
- Part IV. Others
 - ▶ Module 9 (Model display)
 - ▶ Module 10 (Book review)
- Part V. Your work
 - ▶ Module 11 (Preliminary presentation)
 - ▶ Module 12 (Final presentation)

```
"Data Visualization"
```

```
## [1] "Data Visualization"
```