



## Syllabus (2021-2)

Course Title	The Foundation of R Programming for Social Problem Solving	Course No.	11367(01)
Department/ Major	Department of Education	Credit/Hours	3
Class Time/ Classroom	This is the online course. Students can watch lecture video and practice course materials anytime they want during the week.		
Instructor	Name: Sunbok Lee	Department: Department of Education	
	E-mail: sunboklee@ewha.ac.kr	Phone: 02-3277-3338	
Office Hours/ Office Location	Jinseonmi-gwan 239 (by appointment)		

## I. Course Overview

## 1. Course Description

Programming skills are becoming increasingly important in many disciplines, and quickly turning into the core competency for undergraduate students. A solid foundation in programming skills will greatly increase your chance of success in academia and industry. This course is designed to help students in liberal arts and social science 1) build a solid foundation of R programming and 2) develop a capability for solving real-world problems using R in creative and interdisciplinary ways. To that end, this course covers various topics that can be categorized into the foundation and application parts. In the foundation part, students build a solid foundation of R programming by learning data structure (vector, factor, list, data frame, array), control flow (if, for, while statement), functions, data visualization using the ggplot2 package, and data wrangling using the tidyverse package. In the application part, students develop a capability for solving real-world problems by learning the basics of machine learning using R. Throughout the course, the focus will be given to developing students capability for solving problems in our society.

## 2. Prerequisites

No prerequisites are required. This course is designed for students with no programming background in liberal arts and social science who want to learn R programming.

## 3. Course Format

Lecture	Discussion/Presentation	Experiment/Practicum	Field Study	Other
100%	%	%		%

This course will be delivered through online in English



#### 4. Course Objectives

By taking this course, students are expected to learn the following learning objectives.

- Students understand R data structure (vector, factor, array, data frame, list) and can choose the best data structure for a given problem.
- Students understand R functions and can define their own functions.
- Students understand R control flow (if, for, while statements) and use them in their code.
- Students can visualize their data using the ggplot2 package in R.
- Students can wrangle their data using the dplyr and tidyr packages in R.
- Students understand the difference between machine learning and traditional statistics.
- Students understand the representative algorithms in machine learning (e.g., random forest, k-nearest neighborhood, ensemble).
- Students can train their own predictive modeling using R.
- Students can collect and handle text data for text mining.
- Students understand the representative algorithms in text mining (e.g. sentiment analysis).

#### 5. Evaluation System

Students will be evaluated based on the **absolute evaluation system**.

Midterm Exam	Final Exam	Quizzes	Presentation	Final Projects	Assignments	Participation	Other
%	%	40%	%	20%	30%	10%	%

##### Quizzes

- Quizzes will **check students' basic understanding** on the course materials. Students will take Quiz on Cyber Campus.

##### Assignments

- Assignments will check **students' ability to apply their understanding** on course materials in real problems. Students will submit their assignment through Turnitin on Cyber Campus.

##### Final Projects

- Students will 1) select a social problem that they are interested in, 2) get data (usually from government's open data <https://www.data.go.kr/>), 3) suggest solutions on that problem based on the results from THEIR OWN data analysis using R.
  - GROUP PROJECT WILL BE ALLOWED, NO MORE THAN 5
  - Examples from previous open data competitions will be very useful resources as samples (지난 각종 공공데이터 경진대회 수상작들은 기말과제를 위한 좋은 예시가 될것임)
  - (ex: [https://bigdata.seoul.go.kr/noti/selectPageListTabNoti.do?r\\_id=P260&bbs\\_seq=&ac\\_type=A1&sch\\_type=&sch\\_text=&currentPage=1](https://bigdata.seoul.go.kr/noti/selectPageListTabNoti.do?r_id=P260&bbs_seq=&ac_type=A1&sch_type=&sch_text=&currentPage=1) )



## II. Course Materials and Additional Readings

### 1. Required Materials

#### Textbook for R (RDS)

R for Data Science. by Hadley Wickham, Garrett Grolemund. Released December 2016.  
Publisher(s): O'Reilly Media, Inc

※ **You DON'T NEED to buy this book** because the author of this book (Hadley Wickham) kindly released this wonderful book FOR FREE. You can freely read the entire book online here: <https://r4ds.had.co.nz/>

※ (Optional) You can buy Korean version of this book. However, online version should be more up-to-date:

R을 활용한 데이터 과학, 해들리 위컴, 개럿 그롤문드 지음, 김설기, 최혜민 옮김, 인사이트

### 2. Supplementary Materials

# RStudio provides Cheatsheets which summarize R packages in one or two pages:  
<https://rstudio.com/resources/cheatsheets/>

These cheatsheets will be very useful for students to review the functionality of each R package.

# Various open data resources

- ex) open data visualization (공공데이터 시각화)

<https://www.data.go.kr/tcs/vas/selectVisualizationListView.do>

- ex) examples of open data utilization (공공데이터 활용사례)

<https://www.data.go.kr/tcs/puc/selectPublicUseCaseListView.do>

- ex) COMPAS (데이터기반 도시문제 해결 플랫폼)

<https://compas.lh.or.kr/>



## III. Course Schedule (15 credit hours must be completed.)

Week	Date		Topics & Class Materials, Assignments
Week 1		Title	Introduction to data analysis for solving problems in our society
		Topics	types of data analysis, examples of problems, open data challenges
		Reading	Leek, J. T., & Peng, R. D. (2015). What is the question?. Science, 347(6228), 1314-1315. (a very short 2-pages Science article talking about the types of data analysis)
Week 2		Title	Introduction to R
		Topics	Introduction to R, RStudio, R markdown, Base-R, tidyverse packages
		Reading	RDS Ch1. Introduction
Week 3		Title	R markdown for Communication
		Topics	Introduction to R markdown, Publishing your own contents on the web
		Reading	RDS Ch27. R Markdown, Ch29. R Markdown formats, Ch30. R Markdown workflow
Week 4		Title	Data Visualization using the ggplot2 package I
		Topics	grammar of graphics for ggplot2, plots for data visualization
		Reading	RDS Ch3. Data visualization
Week 5		Title	Data Visualization using the ggplot2 package II
		Topics	principles of data visualization, annotating plots
		Reading	RDS Ch28. Graphics for communication
Week 6		Title	Data Visualization in Practice
		Topics	introduction to various plots, map visualization
		Reading	TBA
Week 7		Title	Base-R
		Topics	data type, data structure, function, if, for loop
		Reading	RDS Ch19. Functions, Ch20. Vectors
Week 8		Title	The dplyr package in the tidyverse
		Topics	dplyr for data transformation
		Reading	RDS Ch5. Data transformation
Week 9		Title	The tidyr package in the tidyverse
		Topics	tidyr for data tidying
		Reading	RDS Ch12. Tidy data
Week 10		Title	The purrr package in the tidyverse
		Topics	purrr for functional programming
		Reading	RDS Ch21. Iteration
Week 11		Title	The stringr package in the tidyverse
		Topics	regular expression, wordcloud
		Reading	RDS Ch14. Strings



Week	Date		Topics & Class Materials, Assignments
Week 12		Title	Exploratory Data Analysis I
		Topics	EDA for univariate and multivariate
		Reading	RDS Ch7. Exploratory Data Analysis
Week 13		Title	Exploratory Data Analysis II
		Topics	EDA for univariate and multivariate
		Reading	TBA
Week 14		Title	Predictive Modeling I
		Topics	Introduction to Machine Learning
		Reading	(Kuhn is the author of the caret package, the R package for predictive modeling) Kuhn, M. (2008). Building predictive models in R using the caret package. Journal of statistical software, 28(1), 1-26.
Week 15		Title	Predictive Modeling II
		Topics	Predictive modeling using the caret package
		Reading	Kuhn's webinar on caret package <a href="https://www.youtube.com/watch?v=7Jbb2ltbTC4">https://www.youtube.com/watch?v=7Jbb2ltbTC4</a>

#### IV. Special Accommodations

\* According to the University regulation section #57-3, students with disabilities can request for special accommodations related to attendance, lectures, assignments, or tests by contacting the course professor at the beginning of semester. Based on the nature of the students' request, students can receive support for such accommodations from the course professor or from the Support Center for Students with Disabilities (SCSD). Please refer to the below examples of the types of support available in the lectures, assignments, and evaluations.

Lecture	Assignments	Evaluation
<ul style="list-style-type: none"> <li>Visual impairment : braille, enlarged reading materials</li> <li>Hearing impairment : note-taking assistant</li> <li>Physical impairment : access to classroom, note-taking assistant</li> </ul>	Extra days for submission, alternative assignments	<ul style="list-style-type: none"> <li>Visual impairment : braille examination paper, examination with voice support, longer examination hours, note-taking assistant</li> <li>Hearing impairment : written examination instead of oral</li> <li>Physical impairment : longer examination hours, note-taking assistant</li> </ul>

– Actual support may vary depending on the course.

☞

\* The contents of this syllabus are not final—they may be updated.