

## **Explanatory Data Analysis (47761)**

### Course Description

This class pursuits to understand a series of tasks for data analysis including data collection, processing, analysis, visualization, management of data operating system.

### Lecture

2 sessions/week

### Prerequisite subject

Introduction of Statistics, R programming

### Text book:

- The class materials (<http://ranking.uos.ac.kr>)
- Hadley Wickham, ggplot2: Elegant Graphics for Data Analysis (Use R!), 2nd ed., Springer.

Supplementary material

- Norman Matloff, The Art of R Programming: A Tour of Statistical Software Design., the 1st ed.

### Contents

- Basic R programming (only this semester)
- Writing math symbols with using Tex (ShareTex)
- Version control and Git
- R graphics
- Explanatory data analysis
- Distribution and likelihood
- Text scrapping in web
- ggplot2 in R
- Visualization of statistics on map

### Grading

Attendance (30%), Midterm (30%), Final (30%) , Homework (10%)

**Important announcement:** This class is originally designed for the students who are able to understand the Basic R programming. But the curriculum is temporally modified in this semester because the class of R programming was not open last year. So, some topic can be skipped.

## Topics

Week	Topic
#1	Introduction Install R and other application Git and Github
#2	Data object in R: part I
#3	Data object in R: part II
#4	R programming
#5	R graphics tool
#6	Explanatory Data Analysis
#7	Explanatory Data Analysis / R color
#8	Midterm
#9	Data preprocessing
#10	Text analytics
#11	Web scrapping
#12	ggplot2
#13	Map in R
#14	System and R
# 15	Final exam