

Multivariate Statistical Analysis

2019 Spring

School of Industrial Management Engineering, Korea University

1. Course Description

- This module aims to provide students with the theoretical and practical knowledge and skills to obtain, modify, and analyze a large amount of data from various sources.
- Topics covered in this module include basic association rules, supervised learning algorithms (classification & regression) such as k-nearest neighbor, decision tree, artificial neural networks, and unsupervised learning algorithms (clustering) such as K-Means clustering and hierarchical clustering.
- This module comprises lectures and lab exercises with R to develop the practical skills.
- Students must have knowledge about R language (prerequisite: Programming Language for Data Analysis, 2nd semester for the first grade)

2. Lecturer

- Pilsung Kang, Innovation Hall 801A, 02-3290-3383, pilsung_kang@korea.ac.kr
- Course homepage: <https://github.com/pilsung-kang/multivariate-data-analysis>

3. Time, Place, and Textbook

- Time: Tue 14: 00~15:15 (75 min.), Thu 14:00~15:15 (75 min.)
- Place:
 1. New Engineering Hall 224 (산업경영공학부 전산실)
- Textbook: No single textbook is required. Lecture slides and associated materials (R script, data sets, etc.) will be provided through the blackboard and course homepage.

4. Lecture Module

- 2~3 lecture units consists one module: 1~2 unit: classroom lecture(s), 1 unit: R exercise

5. Introduce Yourself

- ✓ Submit your self-introduction slide (max. 5 pages) to the lecturer via E-mail by the end of the 2nd week.
- ✓ Required information: Name, department, e-mail, cell phone number, recent photo(s)

6. Assessments

- 2 exams (midterm exam 20%, final exam 30%, total 50%): 2 pages of cheating papers are allowed
- 4 assignments (40%)
- Attendance (10%)

7. Schedule

Week	Topics
Mar. 05/07	Orientation
Mar. 12	Association Rule Mining: Classroom lecture 1
Mar. 14	Association Rule Mining: Classroom lecture 2
Mar. 19	Association Rule Mining: R exercise
Mar. 21	Clustering: Classroom lecture 1
Mar. 26	Clustering: Classroom lecture 2
Mar. 28	Clustering: R exercise
Apr. 2	Multiple Linear Regression: Classroom lecture 1
Apr. 4	Multiple Linear Regression: Classroom lecture 2
Apr. 9	Multiple Linear Regression: R exercise
Apr. 11	No class (IE conference)
Apr. 16	Logistic Regression: Classroom lecture 1
Apr. 18	Logistic Regression: Classroom lecture 2
Apr. 23	Midterm Exam
Apr. 25	Logistic Regression: R exercise
Apr. 30	Dimensionality reduction: Classroom lecture 1
May. 02	Dimensionality reduction: Classroom lecture 2
May. 07	Dimensionality reduction: R exercise
May. 09	Decision Tree: Classroom lecture 1
May. 14	Decision Tree: Classroom lecture 2
May. 16	Artificial Neural Network: Classroom lecture 1
May. 21	Artificial Neural Network: Classroom lecture 2
May. 23	Artificial Neural Network: Classroom lecture 3
May. 28	Artificial Neural Network: R exercise
May. 30	Ensemble: Classroom lecture 1
Jun. 04	Ensemble: Classroom lecture 2
Jun. 06	No class (National holiday)
Jun. 11	Ensemble: Classroom lecture 3
Jun. 13	Ensemble: R exercise
Jun. 18	Final Exam