

Introduction to Data Science

Winter, 2019-2020

Instructor: Ick Hoon Jin

1. Basic Information

- Location:
- Time: 09:30 - 12:00
- Instructor: Dr. Ick Hoon Jin
 - Office: 421 Daewoo Hall
 - E-mail:
 - Office Hour:

2. Course Description

This is an introductory course in data science intended for students in a wide variety of areas of study. Students will learn data science using R (one of the well-known statistical softwares in the world). Topics include the basic R programming skills, data visualization, statistics with R, and machine learning.

3. Prerequisite

- Basic Calculus Course or Equivalent

4. Class

- Review: 30 min
- Lecture: 120 min

5. Grading Policy

- 10 Small Assignments: 30%
- 3 Big Assignments: 30%
- Final Exam: 40%
- Total Grade: 10 Small Assignments (3%×10) + 3 Big Assignments (10%×3) + Final Exam

6. Textbook

- Rafael A. Irizarry (2019). Data Analysis and Prediction Algorithms with R.
<https://rafalab.github.io/dsbook/>

7. Topics

- Dec. 27: R and Programming Basics
- Dec. 30: Handling Datasets
- Dec. 31: Introduction to Data Visualization and ggplots
- Jan. 02: Visualizing Data Distributions
- Jan. 03: Data Visualization in Principles and in Practices (Big HW1)
- Jan. 06: Probability and Random Variables
- Jan. 07: Statistical Inferences
- Jan. 08: Statistical Models
- Jan. 09: Regression
- Jan. 10: Linear Models (Big HW2)
- Jan. 13: Introduction to Machine Learning
- Jan. 14: Smoothing
- Jan. 15: Cross Validation
- Jan. 16: Clustering (Big HW3)
- Jan. 17: Final Exam

8. Instructor's Profile

Dr. Jin received his Ph.D in Statistics at Texas A&M University. After spending three years at the University of Texas MD Anderson Cancer Center as a postdoctoral fellow, he joined the University of Notre Dame as an assistant professor and worked there for four years. He is now an assistant professor in Yonsei university.