Lecture 0
Course Information

DSA 8070 Multivariate Analysis

Whitney Huang
Clemson University

About the Instructor

About the Instructor

- Fourth-year Assistant Professor of Applied Statistics and Data Science

- Born in Laramie, WY, grew up in Taiwan

- Obtained a B.S. in Mechanical Engineering, switched to Statistics in graduate school

- Got a Ph.D. (Statistics) in 2017 at Purdue University.

How to reach me?

- **Email:** wkhuang@clemson.edu

- **Office**: O-221 Martin Hall

- **Office Hours**: TBD. Please fill in your availability at
  https://www.when2meet.com/?20891483-jQKTx

Class Policies

Logistics

- There will be two projects. The due dates are:

    - **Project I**: Oct. 19, Thursday

    - **Project II**: Dec. 14, Thursday

- There will be weekly R Labs:

    - To be uploaded to Canvas by 11:59 pm ET on the due dates

    - Worst grade will be dropped

- No lectures during Thanksgiving week (Nov. 20-24)

- Course syllabus / Announcements

- Lecture slides/notes/videos

- R Labs/Projects

- Data sets for lectures and labs

Reference Books

- *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*, **Alan Izenman, 2008**, [Link]

- *Applied Multivariate Statistics with R,* **Daniel Zelterman, 2015** [Link]

- *Methods of Multivariate Analysis*, 3_{rd} Edition, **Alvin Rencher and William Christensen, 2012** [Link]

- *Applied Multivariate Statistical Methods,* 6_{th} Edition, **Richard Johnson and Dean Wichern, 2008** [Link]

Evaluation

Grades will be weighted as follows:

R Labs	20%
Project I	40%
Project II	40%

Final course grades will be assigned using the following grading scheme:

>= 90.00	A
88.00 ~ 89.99	A-
85.00 ~ 87.99	B+
80.00 ~ 84.99	B
78.00 ~ 79.99	B-
75.00 ~ 77.99	C+
70.00 ~ 74.99	C
68.00 ~ 69.99	C-
<= 67.99	F

Computing

We will use software to perform statistical analyses.
Specifically, we will be using R/Rstudio

- a **free/open-source** programming language for statistical analysis

- available at `https://www.r-project.org/` (R); `https://rstudio.com/` (Rstudio)

Topics

Week	Dates	Topic
1	8/23 - 8/25	Introduction
2	8/28 - 9/1	Characterizing and Displaying Multivariate Data
3	9/4 - 9/8	A Short Review of Matrix Algebra
4	9/11 - 9/15	Multivariate Normal Distribution and Copula
5	9/18 - 9/22	Inferences about a Mean Vector
6	9/25 - 9/29	Comparisons of Several Mean Vectors
7	10/2 - 10/6	Multivariate Linear Regression
8	10/9 - 10/13	Repeated Measures Analysis
9	10/16 - 10/20	Principal Components Analysis
10	10/23 - 10/27	Factor Analysis
11	10/30 - 11/3	Canonical Correlation Analysis
12	11/6 - 11/10	Discrimination and Classification
13	11/13 - 11/17	Cluster Analysis
14	11/20 - 11/24	**No Class–Thanksgiving**
15	11/27 - 12/1	Multidimensional Scaling
16	12/4 - 12/8	Review