Lecture 0

Course Information

DSA 8070 Multivariate Analysis
August 15-19, 2022

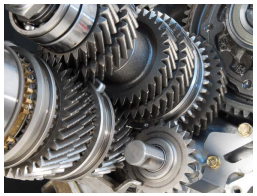
Whitney Huang
Clemson University

About the Instructor

About the Instructor

- Fouth-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[1]
- **Office Hours**: TBD. Please fill in your availability at https://www.when2meet.com/?16411569-cxSc5

[1] I will be visiting the Institute for Mathematical and Statistical Innovation 1155 East 60th Street Chicago, IL from Sept. 13 to Dec. 11

Class Policies

Logistics

- There will be two projects. The due dates are:
  - **Project I**: Oct. 20, Thursday
  - **Project II**: Dec. 15, 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. 22-26)

Course Materials at CANVAS

- 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	8/24 - 8/26	Introduction
2	8/29 - 9/22	Characterizing and Displaying Multivariate Data
3	9/5 - 9/9	A Short Review of Matrix Algebra
4	9/12 - 9/16	Multivariate Normal Distribution and Copula
5	9/19 - 9/23	Inferences about a Mean Vector
6	9/26 - 9/30	Comparisons of Several Mean Vectors
7	10/3 - 10/7	Multivariate Linear Regression
8	10/10 - 10/14	Repeated Measures Analysis
9	10/17 - 10/21	Principal Components Analysis
10	10/24 - 10/28	Factor Analysis
11	10/31 - 11/4	Canonical Correlation Analysis
12	11/7 - 11/11	Discrimination and Classification
13	11/14 - 11/18	Cluster Analysis
14	11/21 - 11/25	**No Class–Thanksgiving**
15	11/28 - 12/2	Multidimensional Scaling
16	12/5 - 12/9	Review