

Lecture 0

Course Overview

DSA 8020 Statistical Methods II
January 4, 2021

Whitney Huang
Clemson University

About the Instructor

Class Policies

Class Overview

About the Instructor

About the Instructor

- **Second year** Assistant Professor of Applied Statistics and Data Science
- Born in Laramie, Wyoming, 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 ([WFH this semester](#))
- **Office Hours:** TBD via Zoom and by appointment

Class Policies

- There will be [three projects](#). The (tentative) due dates are:
 - **Project I:** Feb. 28, Sunday
 - **Project II:** Mar. 28, Sunday
 - **Project III:** Apr. 30, Friday
- 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 in the week Mar. 15-19 ([Spring Break](#))

- [Course syllabus / Announcements](#)
- [Lecture slides/notes/videos](#)
- [R Labs/Projects](#)
- [Data sets for lectures and labs](#)

- *Linear Models with R*, 2nd Edition, Julian J. J. Faraway, 2014 [\[Link\]](#)
- *Extending the Linear Model with R*, 2nd Edition, Julian J. J. Faraway, 2016 [\[Link\]](#)
- *A First Course in Design and Analysis of Experiments*, Gary Oehlert, 2010 [\[Link\]](#)
- *Design and Analysis of Experiments*, 2nd Edition, Angela Dean, Daniel Voss, and Danel Draguljic, 2017 [\[Link\]](#)
- *An Introduction to Statistical Learning: with Applications in R*, Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, 2013 [\[Link\]](#)
- *Time Series Analysis with Applications in R*, 2nd Edition, Jonathan Cryer and Kung-Sik Chan, 2008 [\[Link\]](#)
- *Handbook of Spatial Statistics*, Alan Gelfand, Peter Diggle, Peter Guttorp, andMontserrat Fuentes, 2010 [\[Link\]](#)

Evaluation

Grades will be weighted as follows:

R Labs	25%
Project I	25%
Project II	25%
Project III	25%

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

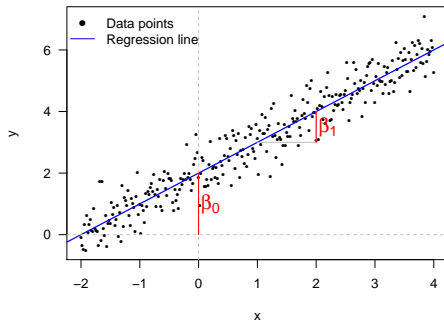
We will use software to perform statistical analyses.

Specifically, we will be using R/Rstudio  /  Studio

- a **free/open-source** programming language for statistical analysis
- available at <https://www.r-project.org/> (R);
<https://rstudio.com/> (Rstudio)

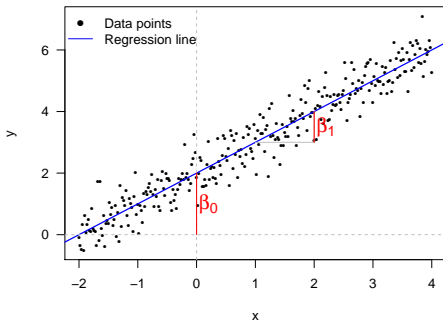
Class Overview

Part I: Regression Analysis



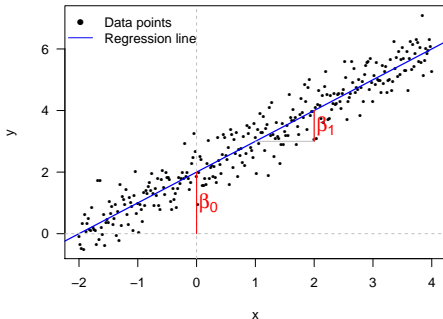
● Multiple Linear Regression

Part I: Regression Analysis



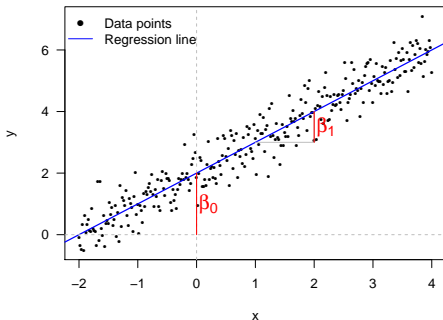
- Multiple Linear Regression
- Regression with Quantitative and Qualitative Predictors

Part I: Regression Analysis



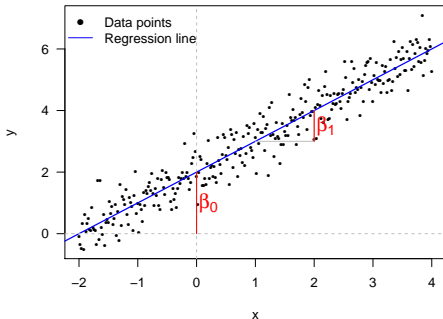
- Multiple Linear Regression
- Regression with Quantitative and Qualitative Predictors
- Nonlinear and Non-parametric Regression

Part I: Regression Analysis



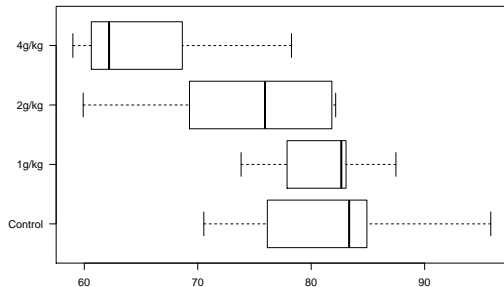
- Multiple Linear Regression
- Regression with Quantitative and Qualitative Predictors
- Nonlinear and Non-parametric Regression
- Ridge Regression and Lasso

Part I: Regression Analysis



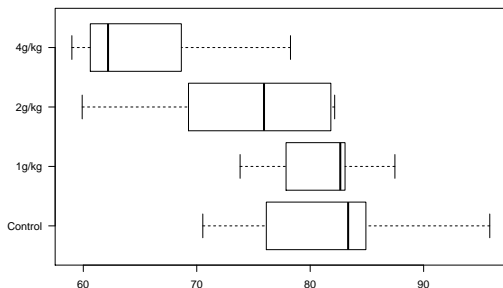
- Multiple Linear Regression
- Regression with Quantitative and Qualitative Predictors
- Nonlinear and Non-parametric Regression
- Ridge Regression and Lasso
- Logistic Regression and Poisson Regression

Part II: Experimental Design



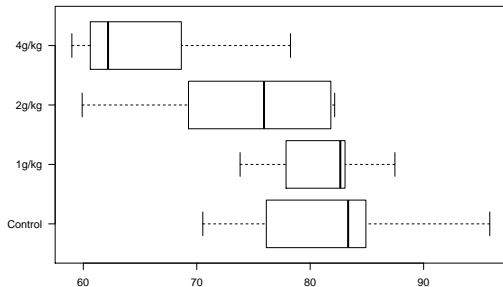
- Introduction to Experimental Design

Part II: Experimental Design



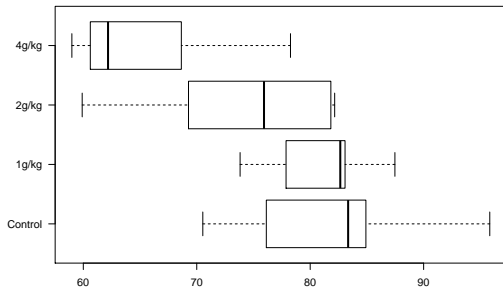
- Introduction to Experimental Design
- Completely randomized Designs, Block Designs, Latin Square Designs, Nested and Split-Plot Designs

Part II: Experimental Design



- Introduction to Experimental Design
- Completely randomized Designs, Block Designs, Latin Square Designs, Nested and Split-Plot Designs
- Random and Mixed Effects Models

Part II: Experimental Design



- Introduction to Experimental Design
- Completely randomized Designs, Block Designs, Latin Square Designs, Nested and Split-Plot Designs
- Random and Mixed Effects Models
- Computer Experiments

- PCA, Classification and Cluster Analysis
- Time Series Models and Stationary Processes
- Interpolation of Spatial Data