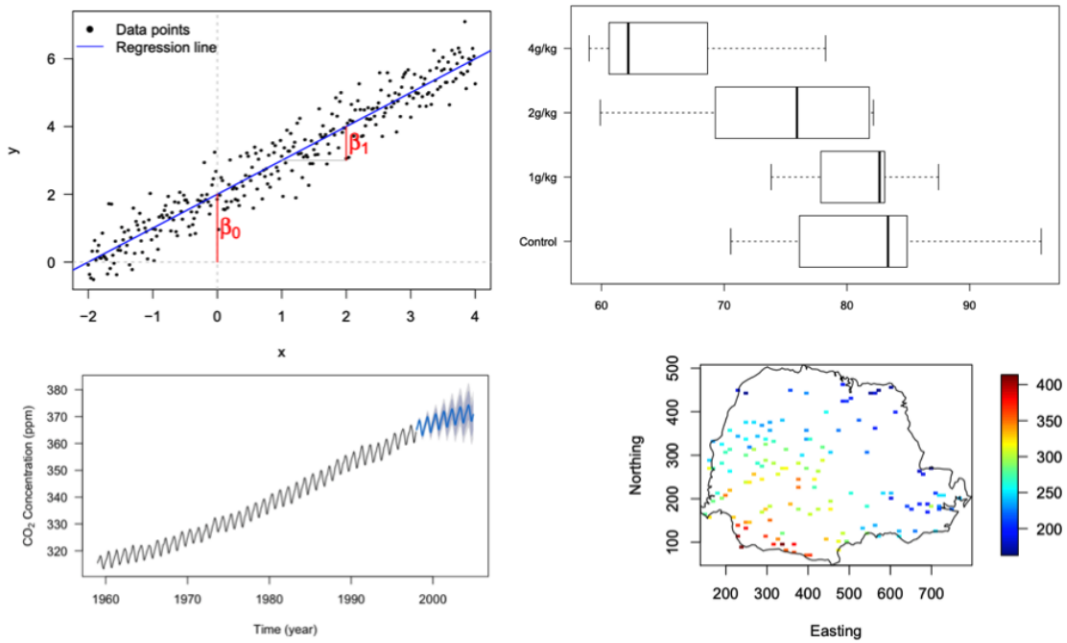
Lecture 0

Course Overview

DSA 8020 Statistical Methods II

Whitney Huang
Clemson University

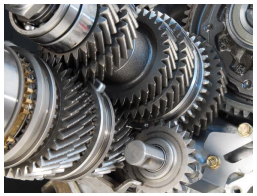


About the Instructor

About the Instructor

- Fifth year Assistant Professor of Statistics
- Born in Laramie, Wyoming, grew up in Taiwan

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



- Ph.D. in Statistics at Purdue, two-year postdoc at Research Triangle, NC and Victoria, Canada

NSF Duke NCSU UNC

How to reach me?

- **Email:** wkhuang@clemson.edu
- **Office**: O-221 Martin Hall
- **Office Hours**: Wednesdays 8:00pm - 9:00pm ET via Zoom and by appointment

Class Policies

Logistics

- There will be three projects. The due dates are:
  - **Project I**: Feb. 22, Thursday
  - **Project II**: Apr. 4, Thursday
  - **Project III**: May 2, 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 in the week Mar. 18-22 (Spring Break)

Course Materials at CANVAS

- Course syllabus / Announcements
- Lecture slides/notes/videos
- R Labs/Projects
- Data sets for lectures and labs

Reference Books

- *An Introduction to Statistical Learning: with Applications in* R*,* 2_{nd} Edition, Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, 2021 [Link]
- *Linear Models with* R*,* 2_{nd} Edition, Julian Faraway, 2014 [Link]
- *Extending the Linear Model with* R*,* 2_{nd} Edition, Julian Faraway, 2016 [Link]
- *A First Course in Design and Analysis of Experiments,* Gary Oehlert, 2010 [Link]
- *Design and Analysis of Experiments,* 2_{nd} Edition, Angela Dean, Daniel Voss, and Danel Draguljic, 2017 [Link]

Evaluation

Grades will be weighted as follows:

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

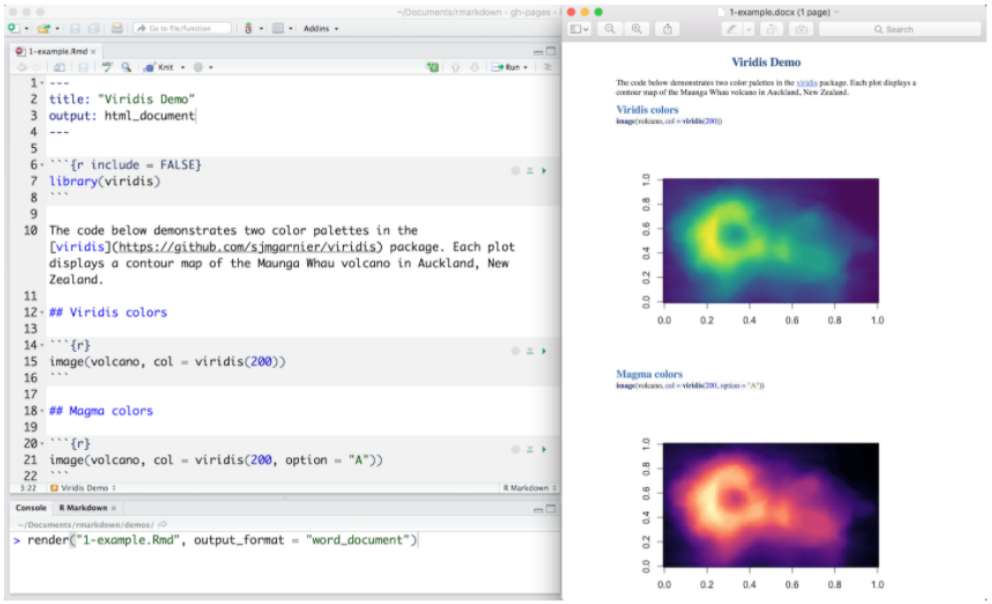
Final course grades will be assigned using the following grading scheme:

>= 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

Computing

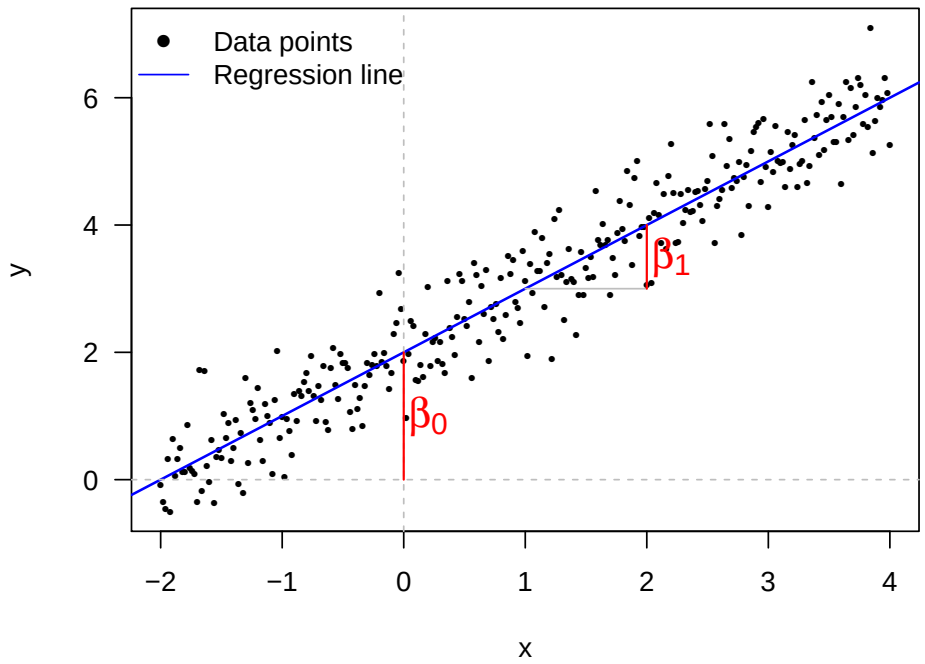
We will use software to perform statistical analyses. Specifically, we will be using `R/Rstudio`

- a **free/open-source** programming language for statistical analysis
- available at `https://www.r-project.org/` (`R`); `https://rstudio.com/` (`Rstudio`)
- We will use **`R` Markdown** for homework assignments



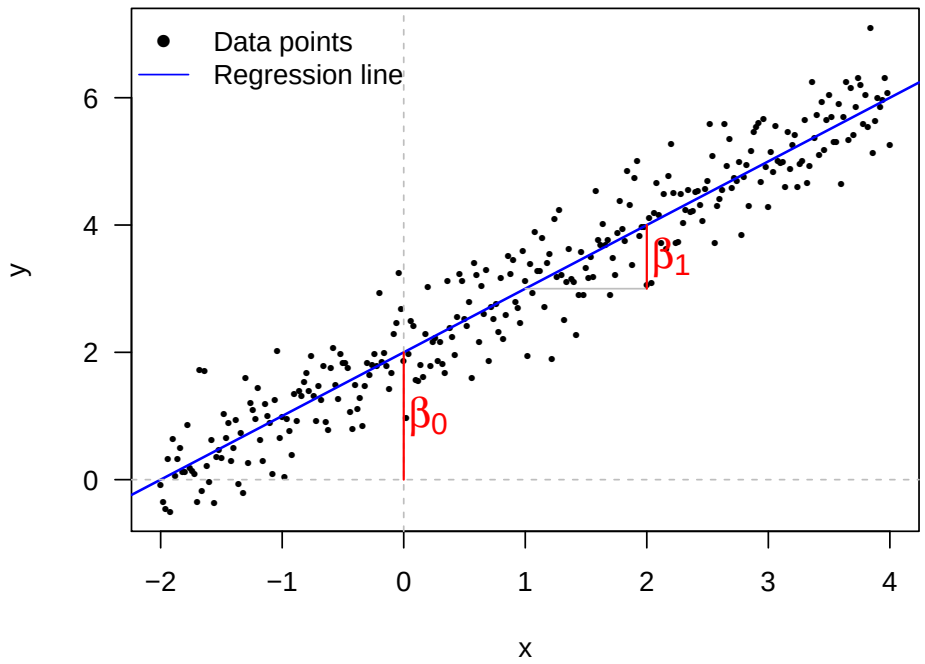
Class Overview

Part I: Regression Analysis (Week 1 - Week 7)



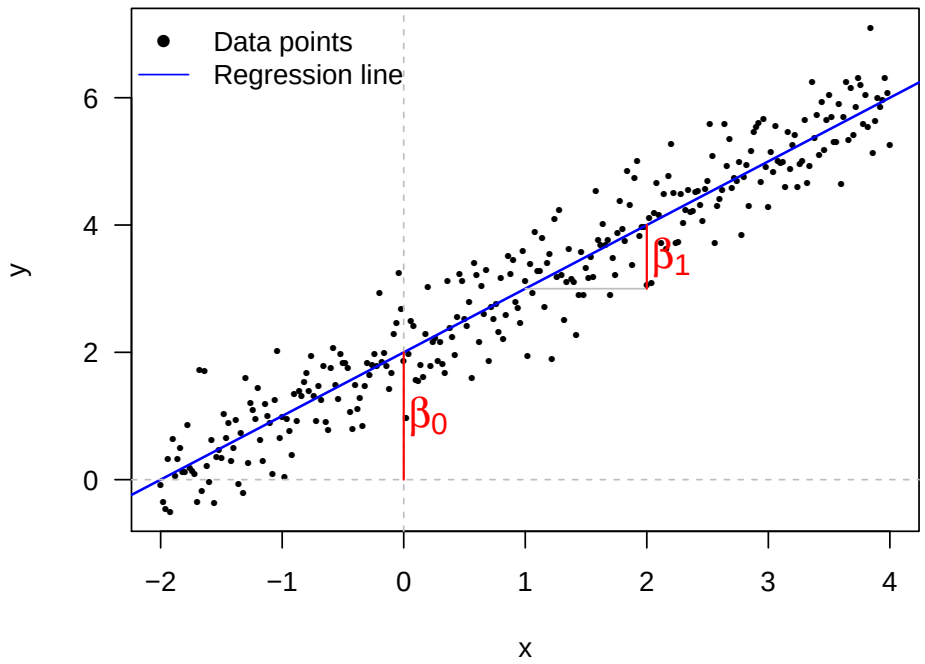
- Multiple Linear Regression

Part I: Regression Analysis (Week 1 - Week 7)



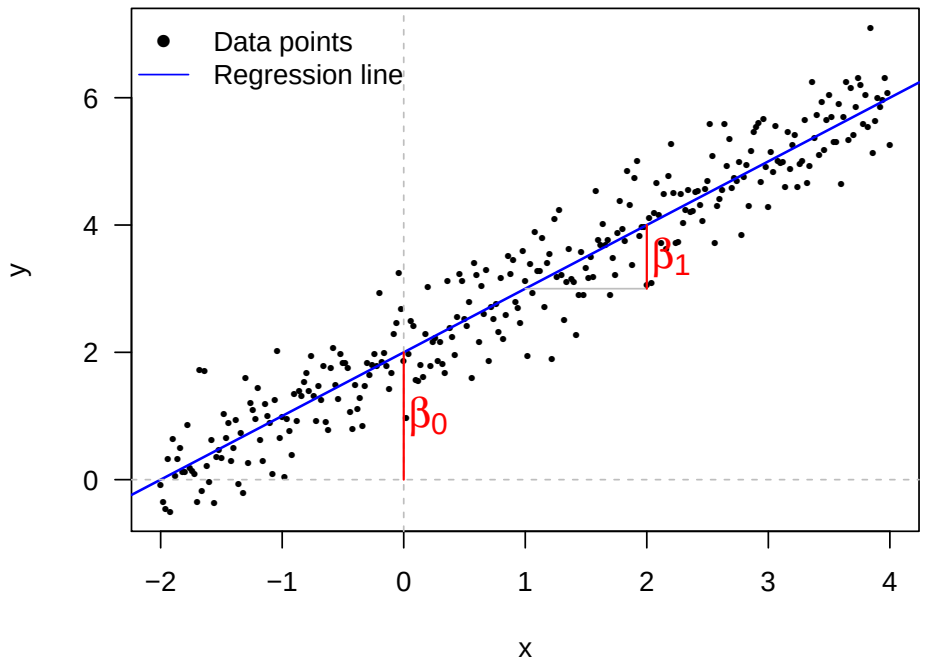
- Multiple Linear Regression
- Regression with Quantitative and Qualitative Predictors

Part I: Regression Analysis (Week 1 - Week 7)



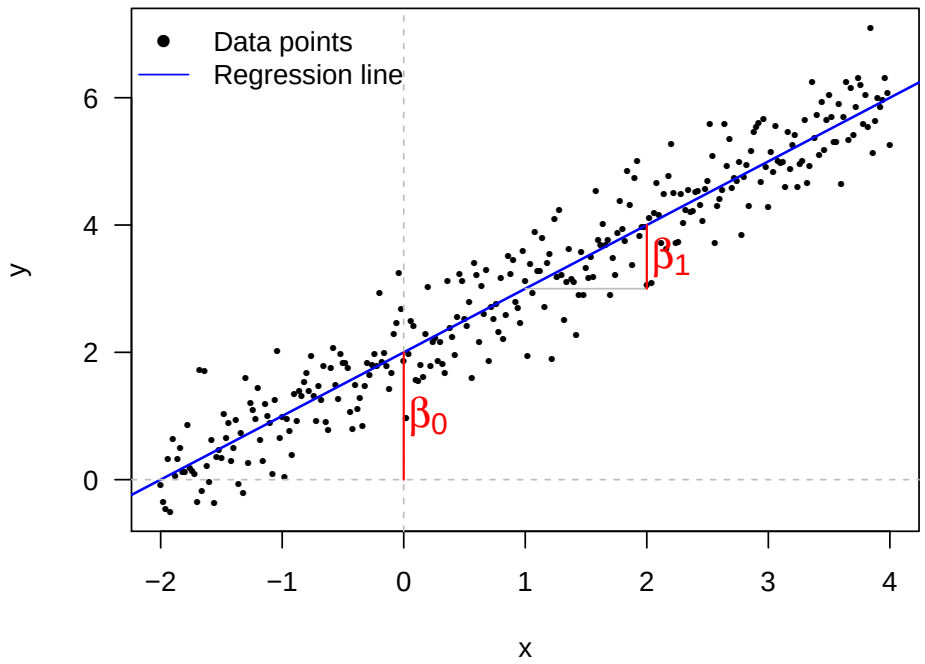
- Multiple Linear Regression
- Regression with Quantitative and Qualitative Predictors
- Nonlinear and Non-parametric Regression

Part I: Regression Analysis (Week 1 - Week 7)



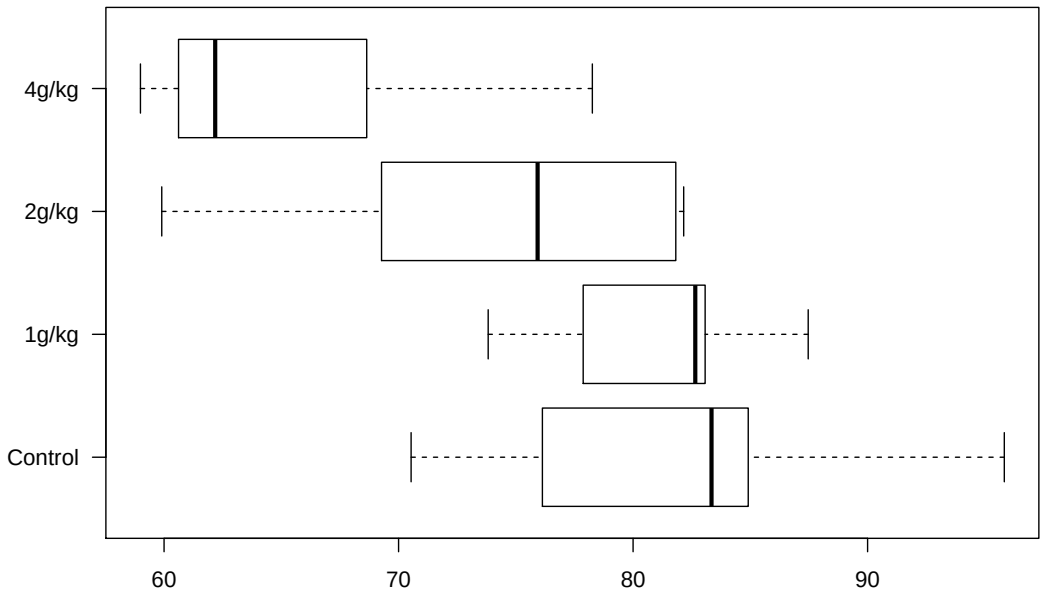
- Multiple Linear Regression
- Regression with Quantitative and Qualitative Predictors
- Nonlinear and Non-parametric Regression
- Ridge Regression and Lasso

Part I: Regression Analysis (Week 1 - Week 7)



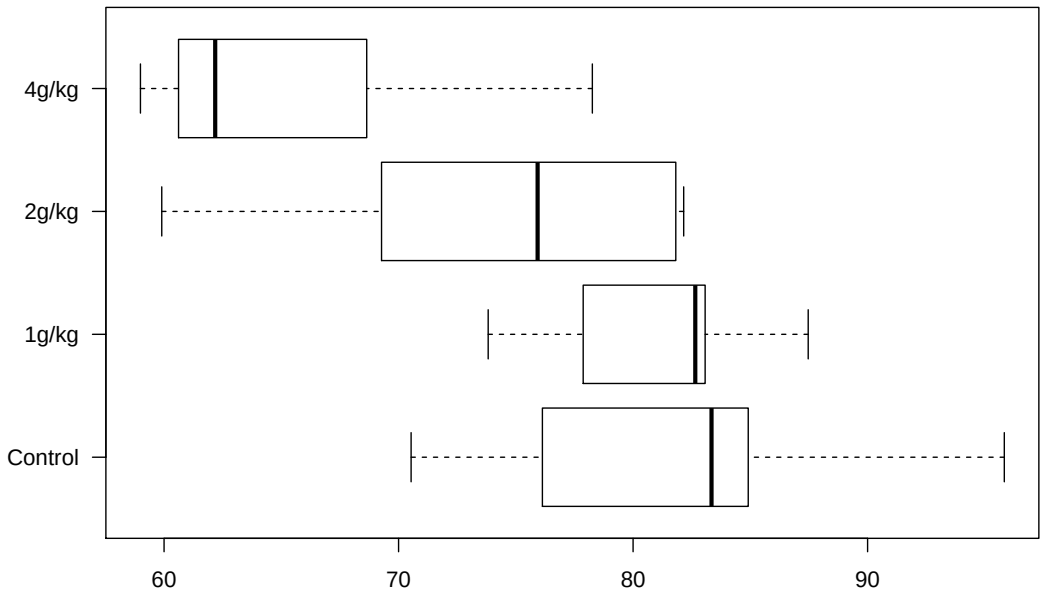
- Multiple Linear Regression
- Regression with Quantitative and Qualitative Predictors
- Nonlinear and Non-parametric Regression
- Ridge Regression and Lasso
- Logistic Regression and Poisson Regression

Part II: Experimental Design (Week 8 - Week 12)



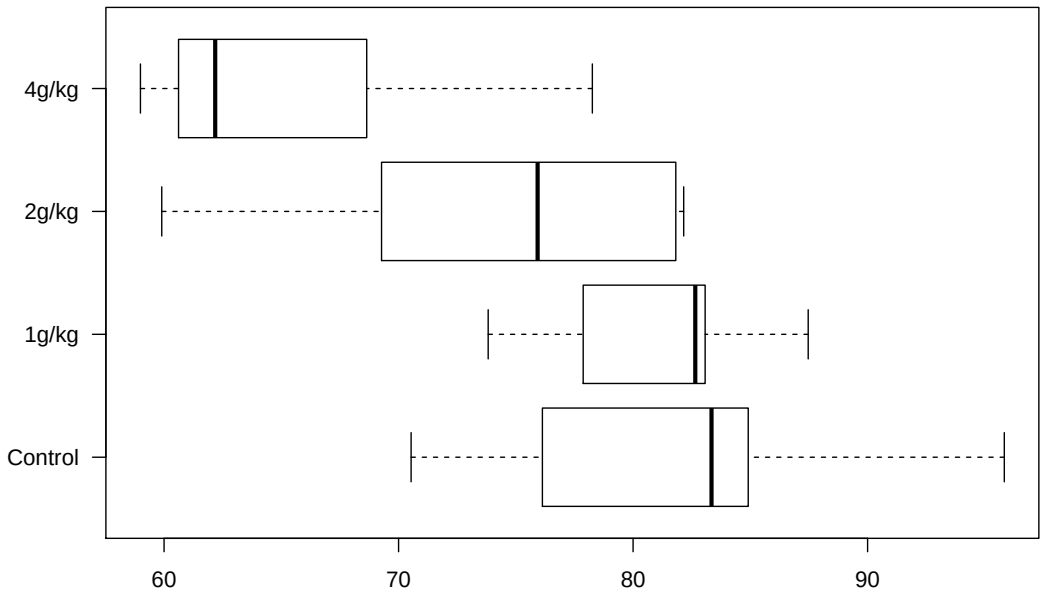
- Introduction to Experimental Design

Part II: Experimental Design (Week 8 - Week 12)



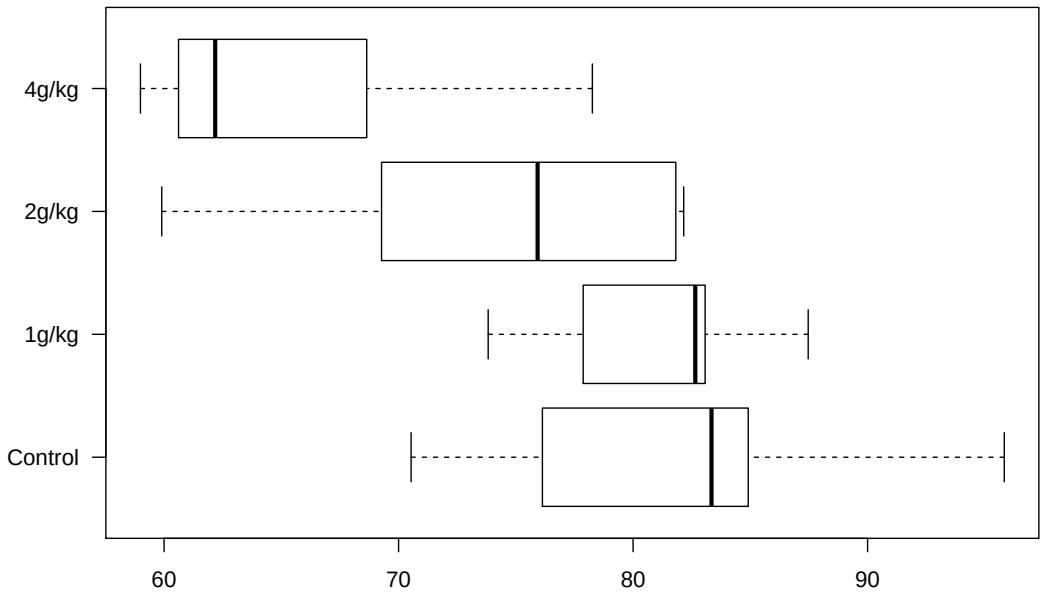
- Introduction to Experimental Design
- Completely Randomized Designs, Block Designs, Nested and Split-Plot Designs

Part II: Experimental Design (Week 8 - Week 12)



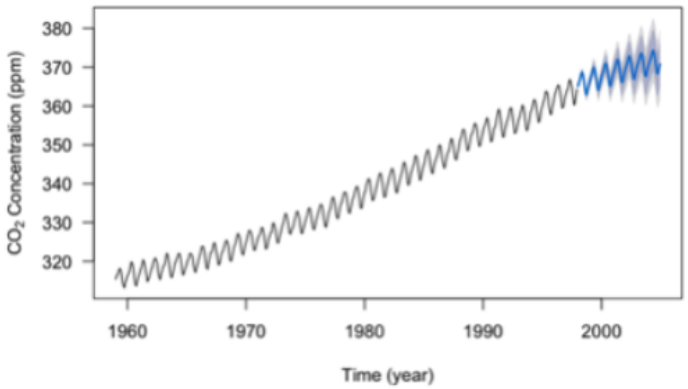
- Introduction to Experimental Design
- Completely Randomized Designs, Block Designs, Nested and Split-Plot Designs
- Random and Mixed Effects Models

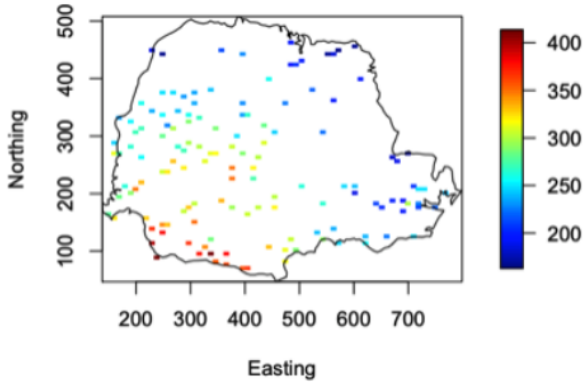
Part II: Experimental Design (Week 8 - Week 12)



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

Part III: Spatial and Time Series Analysis (Week 13 - Week 16)





- Stationary Processes, Autocovariance Function
- Autoregressive Integrated Moving Average (ARIMA) Models and Seasonal Models
- Stationarity and Isotropy, Covariance Function
- Gaussian Process Spatial Interpolation (aka Kriging)