FOS 6936/HUN 6936: Topics in Food Science/Topics in Human Nutrition

Applied data analysis in food science and human nutrition

Class #28928/28927

3 Credit Hour

Spring 2024

Location and Time: **Tuesdays, 12:50 pm to 2:45 pm, AFPL 101**

Thursdays, 11:45 am to 12:35 pm, Weimer 22050

Instructor: Diana H. Taft, Ph.D. (pronouns are she/her)
dianataft@ufl.edu
(352)294-3577

Office Hours: Dr. Taft will remain available for students for 1 hour after each class session in her office (if you plan to come to office hours later during the hour, please let Dr. Taft know by the end of class). Dr. Taft is also available for office hours by appointment, please email to schedule a time.

Course Website:

Required Readings: There is no required textbook for this course. Readings are listed on the syllabus and will be posted on the course website or available through course reserve

Prerequisites: STA 6166 or instructor permission

Purpose of Course: The purpose of this course is deepen student understanding of how to apply statistics to research questions, and to introduce statistical models relevant to food science and human nutrition research not covered in STA 6166.

Course Goals and Objectives: By the end of this course, students will:

1. Be able to choose between basic statistical models to answer research questions
2. Identify when data are a poor fit for a basic model and know what information to bring to a statistician for advice
3. Know when repeated measure or longitudinal data analysis models are required, and be familiar with at least one such model
4. Be able to define machine learning, and be able to use and understand at least one such algorithm
5. Be able to interpret and explain statistical results to non-experts

Grading Policies:

Assignment	Percentage of Final Grade
Homework: Completion	25%
Homework: Correct	25%
Midterm Project	25%
Final Project	25%

Homework: All homework assignments are due by the start of class on the due date. There will be one homework assignment assigned each week and due the following week, except for the weeks the final and midterm projects are due. Homework grades will have two components, completion and correct. Completion points will be obtained by attempting to answer every question on a homework assignment, regardless of the correctness of the answer. Correct points will be obtained for correctly answering the questions. Because some weeks are busier than others, students may opt to not complete 1 homework assignment during the semester, and instead turn in a note stating that this will be the skipped homework, for no grade penalty. A one week extension will be granted on homework assignments if a student informs Dr. Taft that he/she/they are ill as long as Dr. Taft is notified as soon as possible of the illness, and prior to the deadline to turn in the assignment – so if you email at 11 am Thursday saying you are sick, you can have the extension. But if you email at 12 pm (after the deadline), you can't have the extension. One week extensions on homework assignments for other important reasons (i.e. religious holidays, major family events) provided Dr. Taft is notified 1 week before the due date of the need for an extension. Late homework will not be accepted without prior granting of an extension.

Projects: Start working on these early. Dr. Taft is happy to discuss your plans for the project and to provide feedback and assistance, up to one week before the deadline. No extensions will be granted for the final project.

Class Attendance: Dr. Taft does not take attendance. If you need to skip class, as long as your homework is in on time, there is no problem. There are ongoing COVID infections and unusually high levels of influenza and RSV circulating in the community, and some students may be immunocompromised or have small children or vulnerable roommates/family at home. If you are sick DO NOT COME TO CLASS. Dr. Taft is more than happy to review class material with you when you feel better, or to send a zoom link so you do not miss out on the material covered in class. If you skip class every week, do not expect me to spend time catching you up every week, but for the occasional illness/major life event I am happy to be flexible about making sure you get caught up.

Masks: Dr. Taft cannot require you to mask. However, she does work with infants too young to be vaccinated and will provide masks at every class session for those willing to wear one. Please track COVID levels and understand that others in the class may be vulnerable to COVID or interact with those who are more vulnerable.

Schedule

Date	Topic	Required Readings and Homework
January 9, 2024 January 11, 2024	Class Introduction R and RStudio installation Explanation of midterm project requirements	Reading: How to get help in R (https://www.r-project.org/help.html) Introduction to R homework assigned
January 16, 2024 January 19, 2024	January 16: Overview of HiPerGator, Being polite on HiPerGator, How to gain access to HiPerGator	Reading: What is a computational cluster (https://en.wikipedia.org/wiki/Computer_cluster) What is SLURM (https://en.wikipedia.org/wiki/Slurm_Workload_Manager) Introduction to R homework due January 19 HiPerGator homework assigned

	January 19: Tour of HiPerGator	
January 23, 2024 January 25, 2024	Jailbreaking ChatGPT – what it tells us about how AI functions, Overview of AI algorithms, How AI applies to Food Science and Human Nutrition, Ethics in AI for Food Science and Human Nutrition	Reading: https://doi.org/10.1002/hbe2.117 AI and Food Science AI and nutrition HiPerGator homework due AI homework assigned
January 30, 2024 February 1, 2024	Classification and Regression Trees – what are they, how they relate to AI, and what does the computer actually do? How to build CART in R, and what to report if you use CART for a research project	Reading: Background on CART CART in R AI homework due CART homework assigned
February 6, 2024 February 8, 2024	Why probability matters and Bayes theorem (including conditional probability) Introduction to final project Difference between Bayesian and Frequentist statistics (including what those words mean!) and how to figure out what type of statistician you need to talk to	Reading: Dicing with Death Chapter 1 (on course reserve at Marston) CART homework due Probability homework assigned
February 13, 2024 February 15, 2024	Behold the power (Sample Size Calculations, why they matter and how to do them in R and G*Power)	Reading: Statistic review 4: Sample size calculations by Elise Whitely and Jonathan Ball (on Canvas) Probability homework due Reading: pwr package vignette (vignette) Probability homework due Sample size homework assigned
February 20, 2024 February 22, 2024	What is statistical independence anyway? Figuring out the assumptions made by statistical tests, and	Reading: fixed vs random independence assumption https://www.scribbr.com/statistics/anova-in-r/ Sample size homework due ANOVA homework assigned

Commented [TDH1]: Renee – I've already had to cancel this, because there is no way I will have enough students for this to make sense

	why those assumptions matter Introduction to ANOVA	
February 27, 2024 February 29, 2024	Checking your assumptions and linear regression	Reading: ANOVA assumptions linear model assumptions https://en.wikipedia.org/wiki/Linear_regression ANOVA homework due Assumptions homework assigned
March 5, 2024 March 7, 2024	No Class – SPRING BREAK	Enjoy your break!
March 12, 2024 March 14,2024	Midterm project presentations Introduction to non-independent data	Reading: https://doi.org/10.1207/s15328031us0304_2 Midterm project due
March 19, 2024 March 21, 2024	Linear mixed effect models	Reading: Part 1 Part 2 (remember, fixed and random effects have variable meaning. Here the author is using fixed for variables that are constant in a subject and random for variables that can vary within a subject) Assumptions homework due LME homework assigned
March 26, 2024 March 28, 2024	Repeated measures ANOVA and paired t-tests: very useful models in food science and nutrition!	Reading: https://statistics.laerd.com/statistical-guides/repeated-measures-anova-statistical-guide.php LME homework due Repeated measures ANOVA homework assigned
April 2, 2024 April 4, 2024	Moving beyond CART – what is a Random Forest and how to use it	Reading: Random Forest Repeated measures ANOVA homework due Random Forest homework assigned
April 9, 2024 April 11, 2024	What is a statistical analysis plan, and why do you need one? When should you publish your SAP ahead of time, and why?	Reading: https://www.questionpro.com/blog/statistical-analysis-plans/ https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4552232/ Random Forest homework due SAP homework assigned
April 16, 2024 April 18, 2024	Writing up your statistical methods and results – what to include, and what to look for in papers No formal class April 18 – I will be available in the classroom for any students needing extra help with final projects	No reading or additional homework assigned this week – Work on your final project! SAP homework due
April 23, 2024	Final Project Presentations	