GRADTDA 5401: DATA ANALYTICS FOUNDATIONS 1

Instructor Contact Information

Name: **Thomas Metzger**
Office Location: **Cockins 327**
Email: **metzger.181@osu.edu**
Office Hours: **T/R 11-12; R 4:30-5:30;**
By appointment on Zoom

Course Meeting Information

Method: **Online**
Course Hours: ****Course is assumed to contain 40 one-hour lectures and 2 one-hour midterms for a total of 42 contact hours**

Course Prerequisites

Good standing in the MTDA program. Course enrollment is restricted to graduate students enrolled in the program.

Description of Course:

The successful working professional engaged in modern, real world, data science must be able to extract useful information from data and use that information to address work-related challenges. Central to this endeavor is the focus on data-driven decision making under uncertainty. Data come from multiple sources and in all types of formats. Proficiency in identifying, sourcing, manipulating, and interpreting data is paramount.

This is the first course in a two-semester sequence comprised of two, 3 credit-hour courses focusing on R programing and data analysis using R. The sequence is intendent to be taken in parallel with the Big Data Computing Foundations sequence focusing on fundamental CS methods for data science.

Course Learning Goals:

Upon successful completion of the course, students will be able to:

1. Use the statistical programing language and software environment R and the companion integrated development environment RStudio to import, manipulate and visualize data
2. Clean the data and transform them into

formats amenable to statistical analysis and visualization
3. Derive and interpret numerical and visual summaries of the data
4. Conduct application-driven, exploratory analyses that point in the direction of meaningful, application-specific structure in the data, and facilitate the separation of the information "signal" in the data from the "noise" reflecting pure randomness
5. Use their R programing skills to organize the data analysis steps into a streamlined, efficient workflow.
6. Perform basic programing tasks in R
7. Author effective summary reports of the performed analysis using R Markdown

Course Materials and Texts:

Textbook:

Garrett Grolemund and Hadley Wickham (2017), *R for Data Science* (First Edition), O'Reilly; (https://r4ds.had.co.nz/)

Required Software:

- R (www.r-project.org)
- RStudio (www.rstudio.com)

Grading Information:

Grade Breakdown:

Homework = 30%
Projects = 40%
Exams = 30%

Grading Scale:

>90%	A/A-
80-90%	B-/B/B+
70-80%	C-/C/C+
60-70%	D/D+
<60%	E

Course Delivery:

Asynchronous (pre-recorded at student's leisure): each week several lecture videos, totaling approximately 2.5 to 3 hours of lecture, will be posted on the course website. You are responsible for watching the videos and studying the material that is assigned each week. In addition to the lecture videos, weekly assignments will be posted on Carmen.

Synchronous (live): there will also be occasional group discussions or presentations over Zoom, which will be scheduled in advance with students. Office hours and meetings or discussions between individual students and myself can be scheduled in advance and are optional as needed.

Overall, the majority of the course content can be completed on a student's own time at their leisure, within the constraints of deadlines for each lecture video or topic.

Attendance and Participation:

Because this is a distance-education course, your attendance is based on your online activity and participation. Please check your email and log into Carmen at least once per weekday, and participate in discussions on Carmen at least three times per week.

Discussion and Communication:

When posting content visible to myself and/or to other students, please remember to be respectful, thoughtful, and professional.

- Tone and civility: all discussions must contribute to a sense of safety and civility to ensure all students and myself feel comfortable and welcome as part of the course.
- Writing style: please use good grammar, spelling, and punctuation.
- Cite your sources: in academic discussions, please cite your sources to back up what you say. This may be as simple as links to resources you've found online, or references to page numbers in the textbook.

Assignments and Grades:

Homework will account for 30% of your course grade. They will be assigned and submitted via Carmen. Homeworks will usually comprise reflections, data analysis, and coding.

Projects will account for 40% of your course grade. Projects will be completed in small groups and will entail data analysis, coding, and presentation components. Presentations can be given over Zoom at a mutually agreeable time.

There will be a midterm and a final exam that each comprise 15% of your grade. Exams will be completed remotely and individually by students at a predetermined time and submitted on Carmen.

Health and Safety:

The Ohio State University Wexner Medical Center's Coronavirus Outbreak site (https://wexnermedical.osu.edu/features/coronavirus) includes the latest information about COVID-19 as well as guidance for students, faculty and staff. Guidelines and requirements for campus safety from the University's COVID-19 Transition Task Force were published on July 1 on the Safe and Healthy website (https://safeandhealthy.osu.edu). They include the following:

- A daily health check to report body temperature and health status will be required for all faculty, staff and students each day they intend to be on Ohio State's campuses in the autumn.
- Face masks must be worn in indoor settings, including classrooms.
- Members of the campus community will be required to sign a pledge "to affirm their understanding of what is needed to help fight the spread of the virus and their intention to do their part."
- Accountability measures will be in place for those who refuse to abide by required health and safety guidelines.

You are encouraged to complete daily health checks and to self-isolate if running a fever or are in other ways symptomatic.

Potential Disruptions to Education:

If you are unable to complete content or complete assignments because of positive diagnosis or symptoms, please reach out to me as soon as possible to let me know so we can develop a contingency plan. Typically this will include postponing due dates or developing alternative assignments, and will be done on a case by case basis.

If I am unable to create and post new content due to positive diagnosis or symptoms, I will alert the class as quickly as possible to develop a contingency plan.

Lecture Breakdown:

Lecture	Topic
1-3	Introduction and data visualization using R (Textbook Ch 2 and 3) - Loading data - Exploring data - Loading packages - Dataframes - ggplot - correlations
4-6	R coding basics and data transformations (Textbook Ch 4 and 5) - R computations - Variables

	- Functions - Variable types - Logical operators - Data manipulation
7-11	R scripts and introductory elements of exploratory data analysis; Workflow organization and RStudio projects (Textbook Ch 6, 7 and 8) - Exploratory data analysis - Histograms - Boxplots - Outliers - Covariance - Mosaic plots - Barplots - Scatterplots - R Projects
12	Midterm 1
13	Fundamental R data structures: Data frames and Tibbles (Textbook Ch 9-10) - dplyr
14-15	Importing data into R and exporting data to file; Data types (Textbook Ch 11)
16-17	Organizing data into R (Textbook Ch 12)
18-21	Relational data in R (Textbook Ch 13)
22-24	String manipulation is R (Textbook Ch 14)
25-26	Factors in R (Textbook Ch 15)
27	Dates and time (Textbook Ch 16)
28	Midterm 2
29	Programing: Introduction and Pipes (Textbook Ch 17 and 18)
30-31	Programing: Functions (Textbook Ch 19)
32-33	Programing: Vectors (Textbook Ch 20)
34-36	Programing: Iterations (Textbook Ch 21)
37-38	Communication of results: Introduction and R Markdown (Textbook Ch 26 and 27)
39-40	Communication of results: Graphics for communication and R Markdown formats

	(Textbook Ch 28 and 29)

Collaboration and Academic Misconduct:

Students are encouraged to **collaborate** remotely on homework assignments, projects, and discussions, but ultimately the work you submit must be your own. Students should work individually on exams. For all assignments, course and online resources may be used but do not solicit specific help from others, such as by posting specific problems online.

It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "academic misconduct" includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct http://studentlife.osu.edu/csc/.

Disability Services:

The University strives to make all learning experiences as accessible as possible. If you anticipate or experience academic barriers based on your disability (including mental health, chronic or temporary medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion. SLDS contact information: slds@osu.edu; 614-292-3307; slds.osu.edu; 098 Baker Hall, 113 W. 12th Avenue.

Title IX and Sexual Misconduct:

Title IX makes it clear that violence and harassment based on sex and gender are Civil Rights offenses subject to the same kinds of accountability and the same kinds of support applied to offenses against other protected categories (e.g., race). If you or someone you know has been sexually harassed or assaulted, you may find the appropriate resources

at http://titleix.osu.edu or by contacting the Ohio State Title IX Coordinator, Kellie Brennan, at titleix@osu.edu

Diversity and Inclusion:

The Ohio State University affirms the importance and value of diversity in the student body. Our programs and curricula reflect our multicultural society and global economy and seek to provide opportunities for students to learn more about persons who are different from them. We are committed to maintaining a community that recognizes and values the inherent worth and dignity of every person; fosters sensitivity, understanding, and mutual respect among each member of our community; and encourages each individual to strive to reach his or her own potential. Discrimination against any individual based upon protected status, which is defined as age, color, disability, gender identity or expression, national origin, race, religion, sex, sexual orientation, or veteran status, is prohibited.

Mental Health:

As a student you may experience a range of issues that can cause barriers to learning, such as strained relationships, increased anxiety, alcohol/drug problems, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may lead to diminished academic performance or reduce a student's ability to participate in daily activities. The Ohio State University offers services to assist you with addressing these and other concerns you may be experiencing. If you or someone you know are suffering from any of the aforementioned conditions, you can learn more about the broad range of confidential mental health services available on campus via the Office of Student Life's Counseling and Consultation Service (CCS) by visiting ccs.osu.edu or calling 614-292-5766. CCS is located on the 4th Floor of the Younkin Success Center and 10th Floor of Lincoln Tower. You can reach an on call counselor when CCS is closed at 614-292-5766 and 24 hour emergency help is also available through the 24/7 National Suicide Prevention Hotline at 1-800-273-TALK or at suicidepreventionlifeline.org.