

# Stat 4911: Data Analytics Capstone

## Pomerene 160, Spring 2023 TR 28536, WF 33840

Version 2: 1/2/2023

### Course Overview

#### Instructor Information

Instructor: Thomas Metzger, PhD

Email: metzger.181@osu.edu; please do not use my Buckeyemail address.

Office hours: **Virtual (click here)** or in-person meetings (Pomerene 114) by appointment, Tuesdays and Wednesdays, 11 am - 12 pm

Teaching Assistant: Daryl Swartzentruber, swartzentruber.20@osu.edu

#### Course Description

A teamwork-based synthesis of the Data Analytics major curriculum through the analysis of data supplied by a partnering institution. Prepares students for the complexity of data analysis they will encounter outside of the university in a mentored setting. Prereq: 4620, or permission of instructor.

#### Course learning outcomes

By the end of this course, students should successfully be able to:

- Take a general statement of a problem and clearly identify a specific formulation of it that can be studied with available data and information.
- Identify appropriate analytic tools for solving the stated problem.
- Identify shortcomings in an analysis of data and, when possible, use those findings to improve the analysis.
- Draw appropriate conclusions that are supported by an analysis of data.
- Clearly, correctly and effectively communicate methods of analysis and associated findings in writing.
- Clearly, correctly and effectively communicate methods of analysis and associated findings via oral presentation.
- Create visualizations that clearly, correctly and effectively represent data, summarize statistical models and communicate results and findings.

#### Course Delivery and Attendance

This course meets Tuesdays and Thursdays, or, Wednesdays and Fridays in Pomerene 161 from 12:45 pm - 2:35 pm. If it is safe and reasonable for you to do so, class attendance is mandatory. If you cannot attend due to illness, contact tracing, or another reasonable health-related reason, please let me know in advance and follow all of Ohio State's guidelines on health and safety. Each meeting will be labeled with a week number and a letter, "A" (Tuesday or Wednesday meeting) or "B" (Thursday or Friday meeting). For example, the second meeting of week 3 will be Meeting 3B. The last meeting will be Meeting 15B. When practical, meetings will be recorded and posted to Carmen. This is for your benefit to review content and to accommodate students whose health may prevent attendance, but this is not a substitute for in-person attendance.

When it is safe for you to do so, attendance is mandatory and will be taken daily. Most course meetings will consist of group discussions, client meetings and presentations, group work time, and/or statistical

concept review. Each class begins promptly at 12:45 pm, so please be seated, situated, and ready to begin at this time. All absences must be communicated in advance. Excused absences require some indication that you have put the effort in to make up the material and communicate with your team and the clients. Whether an absence is excused is entirely at the discretion of the professor. Health-related absences should be documented with a doctor's note. Strong prior attendance, participation, and advance notice will help your case.

In the event that you are required to miss class due to quarantine, illness, or another mandated reason, **you may join the live class simulcast by clicking here**. Please do not share this link outside of the course.

Participation should include active, respectful engagement in the class by being physically and mentally present, and may also include asking and answering questions, attending office hours, or demonstrating effort and engagement in the course.

## Materials

### Software

- Required software: we will extensively use the statistical software package called R (The R Project for Statistical Computing; <http://www.r-project.org/>). This software package is available for free. You can download R for Windows, Mac, and Linux, from the CRAN archive at <https://cran.r-project.org>. An in-depth introduction to R is available at <http://cran.r-project.org/doc/manuals/R-intro.pdf>. Tutorials are available in the Swirl system, which you can learn about at <http://swirlstats.com/>. "R Programming: The basics of programming in R" is an appropriate first tutorial for students who have never used R.
- Required software: we will also use the R interface RStudio. This package is available for Windows, Mac, and Linux and can be downloaded for free from <http://rstudio.org>. Note that RStudio requires R to be installed.
- Required software: Microsoft Office 365 ProPlus. All Ohio State students are now eligible for free Microsoft Office 365 ProPlus through Microsoft's Student Advantage program. Each student can install Office on five PCs or Macs, five tablets (Windows, iPad® and Android™) and five phones. Students are able to access Word, Excel, PowerPoint, Outlook and other programs, depending on platform. Users will also receive 1 TB of OneDrive for Business storage. Office 365 is installed within your BuckeyeMail account. Full instructions for downloading and installation can be found <https://ocio.osu.edu/kb04733>.
- Optional software: many students may also wish to use other languages such as Python or Ruby, but I am not an expert in these languages and cannot provide the same level of support as I can with R.

## Grades and Assignments

### Readings, Written Work, and Reflections (25%)

Reading assignments will be posted covering topics such as machine learning, data analytics, statistics, data ethics, professionalism, presentation skills, communication, and workplace concerns for data scientists. Reflections may take the form of Carmen discussion boards, live discussions during our class, or written reflections. Submissions will be graded on a combination of the quality, depth, and professionalism of your contribution/reflection. Late written submissions will only be accepted at the professor's discretion. Missed class discussions cannot be made up.

### Projects (75%)

Each project will consist of several weeks of analysis toward a client's needs. Project grades will comprise the following:

- Formal individual project updates. Once per week during our class meetings, each individual member of your group will provide a formal update of their progress. Your formal update should comprise one or two slides with results so far, steps taken, and next steps. It should address how your individual work contributes to the group's overall goals. For this practice project, this will occur just once on Tuesday, January 18, but for our formal projects, this will happen weekly. This is an individual responsibility.
- A single written PDF report, limited to 25 pages, detailing your findings, including the *relevant* code formatted from R Markdown, Jupyter Notebooks, or a similar format. The report should be professional, cleanly formatted, and clear. Interpretations and recommendations should be put into a context relevant to the Port Authority. This is a group responsibility.
- A single presentation of about 15 minutes outlining your findings and recommendations. The presentation is meant to highlight the key takeaways of your analysis, containing recommendations, interpretations, and relevant visualizations that address the key research questions. The audience of your presentation will be decision makers, not statisticians or data scientists! Each team member should contribute with their findings and provide at least one supporting visualization. This is a group responsibility.
- Part of your grade will also come from a peer evaluation, where your teammates honestly and anonymously evaluate your professionalism, contribution, and teamwork. This is an individual responsibility.

## Academic Integrity

I take academic integrity very seriously. There is no place at The Ohio State University for academic misconduct, and I have submitted many students in the past for violations including copying other students' work and receiving unauthorized assistance on exams.

- Submitting work: Collaboration is central to this course and will be for your careers, but ultimately the work you submit must be your own. Directly copying another student's code, work, or file is not permitted.
- Reusing past work: In general, you are prohibited in university courses from turning in work from a past class to your current class, even if you modify it. If you want to build on past research or revisit a topic you've explored in previous courses, please discuss the situation with me.
- Falsifying research or results: All research you will conduct in this course is intended to be a learning experience; you should never feel tempted to make your results or your library research look more successful than it was.
- Collaboration and informal peer-review: The course includes many opportunities for formal collaboration with your classmates. While study groups and peer-review of major written projects is encouraged, remember that copying others' work is not permitted. If you're unsure about a particular situation, please feel free just to ask ahead of time.

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/>.

## Health, Safety, and Accommodations

### COVID Information

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. Be sure to review policies for students that can be found here: <https://safeandhealthy.osu.edu/current-students>.

In the event you are unable to attend class for an extended period of time, please reach out to me as soon as possible. If it is reasonable for you to continue participating remotely, all materials will remain accessible for you. In the event I am unable to participate for an extended period of time, I will let students know as soon as possible. A replacement instructor will be available through the Department of Statistics.

### **Title IX Policy**

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 at [titleix@osu.edu](mailto:titleix@osu.edu).

### **Student Services and Accessibility Accommodations for Students with Disabilities**

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](mailto:slds@osu.edu); 614-292-3307; [slds.osu.edu](http://slds.osu.edu); 098 Baker Hall, 113 W. 12th Avenue.

If you have accommodations, it is your responsibility to communicate these with me in a timely and efficient manner.

Student academic services: <http://advising.osu.edu/welcome.shtml>

Student support services: <http://ssc.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](http://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](http://suicidepreventionlifeline.org).

**Disclaimer**

The materials used in connection with this course may be subject to copyright protection and are only for the use of students officially enrolled in the course for the educational purposes associated with the course. Copyright law must be considered before copying, retaining, or disseminating materials outside of the course.

This syllabus should be taken as a fairly reliable guide for the course content. However, you cannot claim any rights from it and in particular we reserve the right to change due dates or the methods of grading and/or assessment if necessary. Any changes will be communicated to you through official course announcements.