THE OHIO STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

# SYLLABUS: STAT 5301
## INTERMEDIATE DATA ANALYSIS I
## AUTUMN 2020

# Course overview

## Instructor

Instructor: Mario Peruggia
Email address: peruggia@stat.osu.edu
Phone number: 614-292-0963
Class website: https://osu.instructure.com/courses/85839
Lectures: On CarmenZoom, Wednesday and Friday, 8:00–9.50 am (About half of these lectures will be presented asynchronously; see *Course delivery* below)
Office hours: Virtual Hours via CarmenZoom: Wednesday 10:00-11:00 am and Thursday 1:30-2:30 pm. The instructor will also be available to answer any questions via CarmenZoom during class periods scheduled for asynchronous instruction and by appointment.

## Graders

Zilu Liu
Email address: liu.6536@osu.edu

Haozhen Yu
Email address: yu.2823@osu.edu

Tutoring hours will be posted at https://mslc.osu.edu/courses/stat/statistics

## Course description

Statistics 5301 is a first course in a two-semester non-calculus sequence in data analysis covering descriptive statistics, design of experiments, probability, statistical inference, one-sample t, goodness of fit, the two-sample problem, and one-way ANOVA.

**Prerequisites:** The sequence is intended for students with ``limited'' formal mathematics background (a solid grounding in high school algebra is beneficial). However, in terms of data analysis and interpretation, the conceptual level of the course is high. While many of the students in the course are graduate students (it is a required course in many programs), it is certainly an appropriate sequence for junior and senior level undergraduates.

## Course learning outcomes

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

- understand basic concepts of statistics and probability
- comprehend methods needed to analyze and critically evaluate statistical arguments
- recognize the importance of statistical ideas.

## GE Course Information

- This course satisfies the General Education (GE) requirement in **Data Analysis**
- **The GE learning outcomes** for the Data Analysis category are to enable students to deal with problems of data-gathering, presentation, and interpretation. Students should develop an understanding of problems of measurement, be able to deal critically with numerical and graphical arguments, gain an understanding of the impact of statistical ideas in daily life and specific areas of study, and recognize the uses and misuses of statistics and related quantitative arguments. Courses should include exposure to fundamental ideas of probability, involve the use of computer programs in problems of data analysis, and include opportunities to present data using summary measures and graphical techniques.
- The course objectives **address the GE learning outcomes** as follows:

  Students in Statistics 5301 are expected to be able to identify an appropriate analysis for data collected in a study, carry out such an analysis, examine whether the assumptions behind the analysis are reasonable, and recognize the strengths or weaknesses of the study based on how the data were collected. Doing so requires understanding basic concepts in statistics and probability; the ability to create graphical and numerical summaries of data; understanding how the design of a study affects the conclusions that can be made; and the ability to carry out basic statistical analyses (by hand or using statistical software). Students will conduct analyses of data, including a discussion (in plain English) of what conclusions can be drawn.

  The goal of statistics is not calculation but gaining understanding from numbers. This means that the correct numerical answer will only receive partial credit. The remainder of the credit will be available for choosing the best method of solution and explaining

why the method is appropriate. You will also need to interpret your answers in the light of the practical problem.

# Course materials

## Required Textbook

The textbook for this course is being provided via CarmenBooks. Through CarmenBooks, students obtain publisher materials electronically through CarmenCanvas, saving them up to 80%per title. The fee for this material is included as part of tuition and is listed as *CarmenBooks fee* on your Statement of Account. In addition to cost-savings, materials provided through CarmenBooks are available immediately on or before the first day of class. There is no need to wait for financial aid or scholarship money to purchase your textbook.

Unless you choose to opt-out of the program, you do NOT need to purchase any materials for this course at the bookstore. For more information on the program or information on how to opt out, please visit the CarmenBooks website.

[SS] Ramsey F. and Schafer D. (2012), *The Statistical Sleuth: A Course in Methods of Data Analysis*, 3rd Edition, Cengage Learning • ISBN-13:978-1-133-49067-8

Access this eBook through the CarmenBooks reader link in the course navigation.

Note: *The Statistical Sleuth* is also required for Stat 5302.

## Optional Textbook

There is no required textbook for the first half of the course. You may find the following book useful, but it is optional.

[IPS] *Introduction to the Practice of Statistics* (5th Edition onwards) by D.S. Moore and G.P. McCabe

I will highlight other useful references as the course progresses.

# Course technology

For help with your password, university e-mail, Carmen, or any other technology issues, questions, or requests, contact the OSU IT Service Desk. Standard support hours are available at https://ocio.osu.edu/help/hours, and support for urgent issues is available 24x7.

- **Self-Service and Chat support:** http://ocio.osu.edu/selfservice
- **Phone:** 614-688-HELP (4357)
- **Email:** 8help@osu.edu
- **TDD:** 614-688-8743

## Baseline technical skills necessary for online courses

- Basic computer and web-browsing skills
- Navigating Carmen

## Technology skills necessary for this specific course

- CarmenZoom

## Necessary equipment

- Computer: current Mac (OS X) or PC (Windows 10+) with high-speed internet connection
- Webcam: built-in or external webcam, fully installed
- Microphone: built-in laptop or tablet mic or external microphone

## Necessary software

- This class requires you to use the statistical software package called R (The R Project for Statistical Computing; http://www.r-project.org/). This software package is available as Free Software.
    - 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
    - Hands-on tutorials are available in the Swirl system, which you can learn about at http://swirlstats.com/. In particular, "R Programming: The basics of programming in R" is an appropriate first tutorial for students who have never used R.
- An easier to use interface to R is available in the software package 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.**
- More details will be given in live or recorded lectures and on the class web site.
- 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.

## Course delivery

The course will consist of a mix of synchronous and asynchronous content. About half of the class periods (one day a week) will be devoted to asynchronous instruction. The remaining class periods (the other day of the week) will be devoted to synchronous instruction and activities that require direct student participation.

Synchronous content will be presented live over CarmenZoom, and asynchronous content will be delivered by recorded lectures posted on the class website. Short quizzes will be administered each week to test the asynchronous material covered during that week. **Details of the weekly schedule will be announced at the start of each week.**

Each week we will cover approximately 220 minutes of content in total. You will be responsible for watching any live content or recorded videos and studying the material that is assigned. In addition to the lecture videos, assignments will be posted on the class website. You will be given ample time to complete the assignments. The instructor will hold weekly office hours via CarmenZoom according to the schedule given above.

# Grading and faculty response

## Grades

| Assignment or category | Percentage |
|---|---|
| Quizzes | 15 |
| Homework | 15 |
| Midterm 1 | 20 |
| Midterm 2 | 20 |
| Final Exam | 30 |
| Total | 100 |

Grades will be recorded on the class website.

**Homework:** There will be regular homework assignments (about 10). Homework must be uploaded to Carmen by the posted deadline on the day it is due (typically the beginning of class). Homework is **not** accepted by email. Late homework is not accepted, but the lowest homework score will be dropped. You are encouraged to work together on the homework, but

do not copy any part of a homework. Each student must produce his/her own homework to be handed in.

Feel free to ask me for help after you have attempted the questions. The grader for the course does not have the time to provide detailed explanations on each question. To make up for this, I will try to prepare homework solutions detailed enough to allow you to understand how the question could be approached. Homework solutions will be available on the class web site.

**Homework preparation rules:** Homework may be uploaded to Carmen in PDF or Word format. PDF scans of handwritten pages are acceptable. Put your name and the homework assignment number at the top of the first page. Number all pages consecutively. Submit the problems in order, **making sure that the computer output and discussion are placed together** (do not put the computer output at the end of the homework). Include both R code and output in your homework and make it clear what parts of the output are relevant and show how they answer the questions posed in the homework.

**Quizzes:** Short quizzes will be given each week to assess understanding of the material covered asynchronously.

**Exams:** There will be two midterms and one final exam:

| | | |
|---|---|---|
| Midterm 1 | Fri Sep 25 | 8:00–9:50 am |
| Midterm 2 | Fri Nov 6 | 8:00–9:50 am |
| Final | Thu Dec 10 | 8.00– 9.45 am |

All exams are **closed book/closed notes** (except for the sheet of notes described next) and will be **proctored online** – there are no make-up exams. You may use one 8.5x11 inch sheet of paper (both sides), with whatever facts, formulas, or explanations you find helpful, for each exam. Further details will be given in advance of each exam. A basic calculator is allowed – tablets, laptops, cellphones, and other communication devices are not. Statistical tables will be provided as needed.

The first midterm covers the material up to and including Fri Sep 18.
The second midterm covers the material up to and including Fri Oct 30.
The final will cover all the material for the course.

## Faculty feedback and response time

I am providing the following list to give you an idea of my intended availability throughout the course. (Remember that you can call **614-688-HELP** at any time if you have a technical problem.)

### Grading and feedback

For large weekly assignments, you can generally expect feedback within **7-10 days**.

### E-mail

I will reply to e-mails within **24 hours on school days**.

# Attendance, participation, and discussions

Students may miss class, for a variety of reasons related to COVID-19. As much as possible, please stay in contact with the instructor so that we can discuss accommodations should they be needed.

## Student participation requirements

Because this is a distance-education course, your attendance is based on your online activity and participation. The following is a summary of everyone's expected participation:

- **In live lectures:**

  Students will be expected to participate, discuss, and answer questions in online live lectures.

- **Logging in: AT LEAST THREE TIMES PER WEEK**

  Be sure you are logging in to the course in Carmen each week, including weeks with holidays. (During most weeks you will probably log in many times.) If you encounter a situation that might cause you to miss an entire week of class, discuss it with me as soon as possible.

- **Office hours: OPTIONAL OR FLEXIBLE**

  All office hours are optional. If you wish to discuss an assignment with me, please contact me at the beginning of the week if you need a time outside my scheduled office hours.

## Discussion and communication guidelines

The following are my expectations for how we should communicate as a class. Above all, please remember to be respectful and thoughtful.

- **Writing style**: While there is no need to participate in class discussions as if you were writing a research paper, you should remember to write using good grammar, spelling, and punctuation. Informality (including an occasional emoticon) is fine for non-academic topics.
- **Tone and civility**: Let's maintain a supportive learning community where everyone feels safe and where people can disagree amicably. Remember that sarcasm doesn't always come across online.

- **Citing your sources**: When we have academic discussions, please cite your sources to back up what you say. (For the textbook or other course materials, list at least the title and page numbers. For online sources, include a link.)
- **Backing up your work**: Consider composing your academic posts in a text editor or word processor, where you can save your work, and then copying into the Carmen discussion.

# Other course policies

## Health and safety

The Ohio State University Wexner Medical Center's Cornavirus 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).

## Potential disruptions to instruction

- As much as possible, students will have access to material online if they are unable to attend class because of positive diagnosis, symptoms, or quarantine required following contact tracing.
- If the instructor is unable to be present in person because of positive diagnosis, symptoms, or quarantine following contact tracing a new instructor will be assigned to the course. Details will be given on the course website.

## Student academic services

Student academic services offered on the OSU main campus
http://advising.osu.edu/welcome.shtml.

## Student support services

Student support services offered on the OSU main campus http://ssc.osu.edu.

## Academic integrity policy

### Policies for this online course

- **Quizzes and exams**: You must complete the quizzes, midterm, and final exams yourself, without any external help or communication.

- **Written assignments**: Your written assignments, including discussion posts, should be your own original work.
- **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 comparing answers on a quiz or assignment is not permitted. If you're unsure about a particular situation, please feel free just to ask ahead of time.

### Ohio State's academic integrity policy

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/.

## Copyright 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.

## Statement on title IX

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

## Accessibility accommodations for students with disabilities

**The university strives to make all learning experiences as accessible as possible. In light of the current pandemic, students seeking to request COVID-related accommodations may do so through the university's request process, managed by Student Life Disability Services. 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; http://slds.osu.edu; 098 Baker Hall, 113 W. 12th Avenue.**

### Accessibility of course technology

This online course requires use of Carmen (Ohio State's learning management system) and other online communication and multimedia tools. If you need additional services to use these technologies, please request accommodations with your instructor.

- Carmen (Canvas) accessibility
- Streaming audio and video
- Synchronous course tools

## Your 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

## Disclaimer

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.

# Course schedule (tentative)

| Week | Dates | Topics, Readings, Assignments, Deadlines |
|---|---|---|
| 1 | Aug 26, 28 | Data, graphical and numerical summaries, normal distribution.<br>IPS 1.1-1.4 |
| 2 | Sep 2, 4 | Scatterplots, correlation, data sources, design, drawing statistical conclusions<br>IPS 2.1-2.3, 3.1-3.2; SS 1 |
| 3 | Sep 9, 11 | Sampling design, randomness, probability models<br>IPS 3.3, 4.1-4.2 |
| 4 | Sep 16, 18 | Random variables, mean and variance<br>IPS 4.3-4.4<br>Review |
| 5 | Sep 23, 25 | Sampling distribution of a sample mean and of a sampling proportion<br>IPS 5<br>Midterm 1 on Sep 25 |
| 6 | Sep 30, Oct 2 | Confidence Intervals, Hypothesis Testing, Power and Inference as a Decision, Use and Abuse of Tests<br>IPS 6.1-6.4 |
| 7 | Oct 7, 9 | One Sample t-tools, The Sign Test, One-Sided and Two-Sided Tests, (Pooled) Two-Sample t-tools,<br>IPS 7.1-7.2; SS 2 |
| 8 | Oct 14, 16 | Assumptions of the t-tools, Variable transformations<br>SS 3 |
| 9 | Oct 21, 23 | non-pooled t-tools, the Rank Sum Test<br>SS 4 |
| 10 | Oct 28, 30 | Wilcoxon Signed-Rank Test<br>SS 4.4 |

| | | |
|---|---|---|
| | | Review |
| 11 | Nov 4, 6 | Inference for a population proportion, Comparing two proportions<br>IPS 8, SS 18.1-18.2<br>Midterm 2 |
| 12 | Nov 13 | ANOVA Introduction<br>SS 5 |
| 13 | Nov 18, 20 | Inferences about Linear Combinations of Group Means<br>SS 6.1-6.2 |
| 14 | Nov 25 | Simultaneous Inferences<br>SS 6.3-6.4 |
| 15 | Dec 2, 4 | Model Comparison with the F-test<br>Review |