# STS 2300 Brief Overview

## About Me

My name is Dr. Nicholas Bussberg (please feel free to call me Nic), and I am your STS 2300 instructor for Spring 2025. I want to take this opportunity to welcome you to Introduction to Data Analytics! I am a faculty member in the Department of Mathematics & Statistics. I'm originally from Indiana and attended Indiana University for all my schooling. My research is in environmental statistics, but I have many interests outside of the environmental sciences that you'll see throughout the course. This course will teach you the basics of the statistical programming software R, which is what I use in my research. Please don't hesitate to ask me questions about how/why I use R, as there are several approaches to completing tasks in R.

## Course Philosophy

Statistics and coding can be challenging at times. However, without pushing ourselves, we are limited in how much we can learn and improve our skills and understanding (e.g., how a weightlifter won't get stronger if they never lift heavier weights). This means there may be times where you struggle with course material. This is to be expected, and it should not discourage you. Each and every person in this class is fully capable of succeeding. In fact, we will aim to embrace struggle as a productive element of learning and growth. While most people enjoy "getting a problem right", we learn the most when we make mistakes and then work to understand and correct those mistakes. Each mistake is an opportunity to learn and improve, and this *productive struggle* will ultimately make us stronger in the end.

Please reach out me if you have any questions or concerns (nbussberg@elon.edu). I'm more than happy to work with you (see more in the 'Getting Help' section below).

# STS 2300 – Introduction to Data Analytics
## Section A
## (Tentative) Syllabus for Spring 2025

**Instructor:** Dr. Nicholas Bussberg (Nic)

**Email:** nbussberg@elon.edu

**Current Office:** Duke 201A

**My virtual office is:** https://elon.zoom.us/my/nbussberg

**Student Drop-In Hours (held in Duke 201A):**

- o Mondays, 2:00-3:00pm

- o Tuesdays, 11:00am-12:00pm

- o Thursday, 1:00-2:00pm

- o Also by appointment – just email me!

**Class Meetings:**

- o Monday/Wednesday/Friday, 9:30-10:40am, Duke 303

**Teaching Assistant (TA):** Kelly Donovan

**TA Drop-In Hours:**

- o TBD

Course Description

This course will introduce students to a cutting-edge statistical programming language, such as R, and will provide foundational statistical tools for data analytics. Rather than focusing on statistical theory, students will learn to apply computer-intensive randomization-based approaches to statistical inference for a range of scenarios that build upon and expand topics covered in STS 2120. Key data analytics topics will include importing data from a wide variety of sources, data wrangling, data visualization, and exploratory data analysis. Simple and multiple linear regression will be introduced from a predictive modeling perspective. Throughout the course, students will learn to generate reproducible and dynamic statistical reports, with emphasis placed on communicating data analytic results to non-statistical audiences.

## Prerequisite

A student can register for STS 2300 if they have previously taken (or are concurrently registered for) STS 2120 Statistics in Application (or by permission of the Statistics Program Coordinator, Dr. Mark Weaver, mweaver11@elon.edu).

## Course Materials

- Course material (notes, activities, schedule) will be available on Moodle (www.elon.edu/moodle) and/or the course Github repository (https://github.com/nbussberg/STS2300-Spring2025). Come prepared to take notes electronically, on printouts, or on a separate sheet of paper.
- **Primary Textbook**: Statistical Inference via Data Science: A Modern Dive Into R and the Tidyverse by Chester Ismay and Albert Y. Kim (https://moderndive.com/)
- **Additional Text**: R for Data Science by Hadley Wickham, Mine Çetinkaya-Rundel, and Garret Grolemund (https://r4ds.hadley.nz/)
- **Additional Text**: Modern Data Science with R by Benjamin Bauer, Daniel Kaplan, and Nicholas Horton (https://mdsr-book.github.io/mdsr3e/)

## Statistical Software

The course will primarily use the free and open-source statistical coding language R through the IDE RStudio. R and RStudio can be downloaded onto personal computers or accessed on campus computers via AppsAnywhere. There is also a cloud-based version available for $5 per month.

## Getting Help for the Course

Question: Who should get help outside of class?

Answer: Everyone! Each person can benefit from discussing concepts and ideas with others. Nobody gets a special mark on their transcript saying they got a certain grade without seeking help. So, you may as well make use of resources that are available to you.

There are many opportunities to get help outside of class this semester.
- There may be tutors available through the department and through the Belk library. To meet with tutors, you usually need to sign up for a time slot beforehand. Information will be posted on Moodle once available.
- There are open student drop-in hours a few times a week hosted by both the TA and me. You do not need to email to talk with me during these times – simply drop by! If there are many students, we may use the conference table in my office suite.
- Elon's Data Nexus has weekly data technology help sessions. If you need help with technology (e.g., R, Excel), the Data Nexus Technology Ambassadors are a good resource to utilize.
- There is a (free) textbook that we will be following, and I strongly recommend that you keep up with the readings. There are two supplementary textbooks that could also be used for reference and extra practice.
- You can also work with other students in your class. This is encouraged, but you must do your own work – copying answers from someone else is considered a violation of Elon's Honor Code. If you have questions about how to work with classmates without violating the Honor Code, please feel free to reach out to me.

You may see these activities as taking "extra time", but you will likely save time in the long run by getting a chance to correct misconceptions and to solidify your own understanding. Sometimes just saying an idea out loud or explaining it to someone else is enough for it to "click" in your own mind.

## Course AI Policy

You should treat Artificial Intelligence (AI) tools, such as ChatGPT, the same as other online resources. As a general rule, you may use AI as a resource as you complete assignments but not to answer the exercises for you. You are ultimately responsible for the work you turn in; it should reflect your understanding of the course content.

In this course, AI can be used only on homework and code for projects (not Objective Assessments, reflections, or project reports/presentations). This includes code, but you must explicitly cite where you obtained the code. Any recycled code that is discovered and is not explicitly cited will be treated as plagiarism. You may use these guidelines for citing AI-generated content.

There are two guiding principles for how you can use AI in this course:
1. *Cognitive dimension*: Working with AI should not reduce your ability to think clearly and independently. If you use AI, it should facilitate, (i.e., not be used to replace), your learning.
2. *Ethical dimension*: If you use AI, you should be transparent about its use and make sure it aligns with academic integrity standards. In this course, this means that you must cite that you used AI, which AI tool you used, and which parts of your work was influenced by AI.

Final note on AI: AI can be used as an effective tool to help you learn material and help you work more efficiently. However, AI does not always produce correct or appropriate solutions, and you must be able to understand the when to trust the results and when to avoid them. **If you have any questions, please reach out to me**, and we can work through them.

This AI policy was based on materials from Dr. Maria Tackett and Dr. Joel Gladd.

## Diversity

Diversity of student experiences and perspectives is essential to the deepening of knowledge in this course. I consider it part of my responsibility as an instructor to address the learning needs of all students in this course. It is also the responsibility of all students to exhibit professional courtesy and respect for all members of the class. The course materials and discussions are intended to be respectful of social identities. Any suggestions that you have about other ways to incorporate diverse experiences in the course materials or discussions are welcome.

I (like many people) am still in the process of learning about diverse perspectives and identities. If something was said in class (by anyone) that made you feel uncomfortable, please talk to me about it.

I also acknowledge that many of the statistical techniques and even the SAS program were discovered or written mostly by white men. There were times that statistics were used to further ideas that we now know are wrong, such as white supremacy. For instance, Ronald Fisher was one of the founding fathers of modern statistics, but he was also well known for his role in promoting eugenics, or scientific racism. I know we will use data and techniques that were discovered by flawed people, and I will do my best to acknowledge this whenever I can.

## Outcomes

This course will enable students to improve their data competencies using leading statistical software for data analytics, such as R. Specific outcomes that should prove valuable include the ability to:
- Import various types of data.
- Wrangle data into a format appropriate for conducting planned analyses.
- Build multi-layered data visualizations that transcend the default visualizations typically provided by statistical software.
- Generate reproducible and dynamic statistical reports and presentations using R Markdown, or a comparable platform.
- Fit simple and multiple linear regression models for the purpose of predicting a quantitative response variable.
- Test statistical hypotheses and generate confidence intervals for a broad range of data measures using randomization- and permutation-based approaches.
- Conceptualize, plan, and conduct a data analytics project, from beginning to end.
- Effectively communicate data analytic results orally, visually, and in writing.

## Learning Objectives

A list of specific objectives follows each area of content.

1. After introduction to statistical software, students will be able to:
   - Describe different elements of statistical programming.
   - Run basic code to manipulate vectors and data frames in the console and using scripts.

2. After data summaries and reproducible markdown files, students will be able to:
   - Summarize data with tables and numeric summaries (including grouped summaries).
   - Create markdown documents that mix text, code, and output in one file.
   - Customize markdown documents to show/hide various elements, add elements like links/images, update automatically with new data, and more.

3. After data wrangling and data management, students will be able to:
   - Read data of various types (.csv, .xlsx, .txt, and more) into statistical software.
   - Subset data by rows and columns as desired.
   - Create new variables to add to a data frame.
   - Join data sets and convert data from long to wide format (and vice versa).

4. After data visualization, students will be able to:
   - Create various graphs like bar graphs, histograms, boxplots, scatterplots, etc.
   - Create extensions of the graphs above that incorporate additional variables into them with color, size, shape, etc.
   - Customize graph themes, labels, legends, and scales to effectively communicate results.

5. After simple and multiple linear regression, students will be able to:
   - Generate and understand regression equations, including summary statistics like the coefficient of determination.

- Interpret intercept and slope values from a regression line in context.
- Use methods like backward selection to choose an appropriate multiple regression model.
- Use technology to make predictions with regression models.

6. After bootstrap sampling and confidence intervals, students will be able to:
- Describe the idea behind a sampling distribution and a bootstrap resampling distribution.
- Describe how sampling variability is used to construct confidence intervals and why confidence intervals are useful.
- Be able to use the idea of bootstrap resampling to create confidence intervals in a wide range of settings (e.g. mean, difference in means, proportion, difference in proportions).
- Be able to interpret confidence intervals in context.

7. After null distribution generation and hypothesis tests, students will be able to:
- Describe the general concept of hypothesis testing and why it is useful.
- Use simulation to create null distributions to which observed data is compared.
- Describe what a p-value is in terms of a null distribution and explain its limitations.
- Use the concept of simulated null distributions to conduct hypothesis tests and be able to reach appropriate conclusions.

## Planned Assessments

There will be a variety of assessments that will allow you to demonstrate that you mastered the material:

- **Objective Assessments (OA):** Online quizzes will be administered via Moodle each week. These quizzes are designed to assess your mastery of the course objectives. While students are encouraged to discuss the course material with classmates in study sessions, these OAs are to be completed individually and should reflect a student's individual understanding of the material. These assessments are a low-stakes way of measuring understanding in the course. Performance on assessments should help students identify areas that they need additional help on. There will be approximately 13 OAs throughout the course. *To master an OA, you must score at least 80% on that quiz.* You will be given two attempts to complete assessments. Due dates for each OA will be assigned in Moodle. OAs will typically be due on Mondays and will cover primarily content from the previous week.

- **Homework (HW):** Homework will involve higher level understanding relative to objective assessments, as you will need to apply your knowledge to new questions. There will likely be 8 homework assignments throughout the course. For each homework, I will provide feedback and give you either an "Accepted/Pass" (AP) or "Revise and Resubmit" (RR). The AP/RR designation is not a grade; it is simply a very broad-scale indicator of whether there are questions that you can revise and improve. You are allowed two resubmissions per homework assignment to incorporate any feedback I give you. Resubmissions are optional, but strongly encouraged. They are the best way to show that you understand what you did incorrectly.

- **Projects:** Projects are the capstone assessments for the course. They will involve synthesizing information learned throughout the semester and will require students to apply knowledge to new, less-structured settings. There will be 2 projects during the course. Each project will have multiple components, including drafts that will be submitted earlier in the process. Students will receive feedback on the drafts that should be incorporated into the final project.

    Early draft or checkpoint components will graded as complete/incomplete. To be marked as complete, you must answer each question or prompt fully and thoughtfully. The final product for each project will be receive an "Accepted/Pass" (AP) or "Not Pass" (NP) mark.

    Because each project will likely be group-based, each project will also have a companion Individual Assessment to assess your personal understanding of the project and corresponding class concepts. This will also be marked "Accepted/Pass" (AP) or "Not Pass" (NP).


- **Reflections:** During the semester, students will be asked periodically to reflect on their learning progress. The reflections are an opportunity to explain what you have learned, what is important to you, and what you still believe you can work on. Reflections will be graded as complete/incomplete. To be marked as complete, you must answer each question fully and thoughtfully.


## Requirements and Grades

For this course, we will be using a mix of non-traditional grading structures called "ungrading" and specifications grading. This will likely be different from most other courses you have taken. Instead of each assessment counting toward a percentage of your grade, your grade will instead be determined on what you complete during the class. Note that for the above assessments you can either master it (with a grade of 80% or higher) or receive a passing/complete grade.

Your **initial grade range** will be determined based on the criteria outlined in the table on the next page. Note: it is possible to be in different grade ranges for different categories – this will be resolved in our final grade meeting.

Your **final grade** (including assigning a "+" or "-'' as appropriate) will be determined during a final, individual grade discussion with me. Before our meeting, you will complete an end-of-semester reflection that explains, among other topics, what you learned in the class and what final grade you believe you should receive. We will then discuss your responses and settle on a final grade during the meeting. Note: although you will help me evaluate your progress, I reserve the right to modify the final grade based on my observations.

We will discuss the grading policy on the first day of class, during which I am happy to answer any questions you have. We will also have a mid-semester individual meeting, during which we'll discuss where you are at in the course, how well you are doing, and what you should do to finish the course.

I understand that this may cause some uncertainty or anxiety in some of you. You should know that it's ok! At any point during the course, if you are feeling uncertain or anxious, please come talk to me. I'm more than happy to discuss where you are at.

| Grade Range | Criteria |
|---|---|
| **D** | - Show mastery (a score of 80% or higher) on 5 Objective Assessments <u>and</u> attempt at least four others.<br>- Receive a Pass on 4 Homework Assignments <u>and</u> attempt at least two others.<br>- Not Pass either project <u>but attempted</u> in both.<br>- Not Pass either project Individual Assessment <u>but attempted</u> both.<br>- Completes Reflection 1 and 2, but not the final reflection. |
| **C** | - Show mastery on 7 Objective Assessments <u>and</u> attempt at least four others.<br>- Receive a Pass on 5 Homework Assignments <u>and</u> attempt at least two others.<br>- Pass one project <u>and</u> attempted in the other.<br>- Pass one project Individual Assessment <u>and</u> attempted the other.<br>- Completes the final grade reflection. |
| **B** | - Show mastery on 9 Objective Assessments <u>and</u> attempt the other four.<br>- Receive a Pass on 6 Homework Assignments <u>and</u> attempt the other two.<br>- Pass both projects.<br>- Pass one project Individual Assessment <u>and</u> attempted the other.<br>- Completes the final grade reflection <u>and</u> one other reflection. |
| **A** | - Show mastery on all 13 Objective Assessments.<br>- Receive a Pass on 7 Homework Assignments <u>and</u> attempt the other one.<br>- Pass both projects.<br>- Pass both project Individual Assessments.<br>- Completes all three reflections. |

Participation, Attendance, and Late Policy

Learning statistics requires a lot of practice. I expect that you engage in opportunities to hone your skills and think about how to apply those skills to your area of interest. Though this is typically done through in-class participation, there are many other methods. For example, attending tutoring sessions or coming to drop-in hours are great ways to engage in the course content.

Class sessions are a vital component of the course where we will introduce concepts and learn together by working through examples and activities. You are expected to attend and actively participate in all classes. This also includes completing out-of-class work on time, e.g., readings, assessments.

Though attendance and late assignments are not directly included in the initial grade range table above, it will be hard to justify a higher grade for the course if you frequently turn work in late ("frequent" = 3 or more late assignments) or miss many classes (7 or more classes). **Having a lot of late work or missing a lot of classes will result in a lower grade for the class.**

For attendance, everyone has the six "free" absences that apply to any reason. This includes both "excused" (e.g., athletic competitions, religious observances) or "unexcused" (e.g., illness, family emergency) absences. This means that you must plan ahead if you know you will be absent several times during the course. If there are special circumstances that will cause you to miss many days of class in a row, such as a serious illness, please let me know and we can make alternative arrangements.

If you are unable to attend class, filled-in lecture notes will be available. I strongly recommend meeting with a classmate to discuss what you missed. After you review the notes and discuss with a classmate, please feel free to meet with me (or a tutor) to cover the concepts you are unsure about.

Finally, extraordinary circumstances may arise for some of you during the semester. You may find yourself facing more than you can handle, making it difficult to focus on this course. First, Elon's Counseling Services provides support for students going through such problems, and I encourage you to seek their help if needed. Second, I try to be flexible with these disruptions and, on a case-by-case basis, may provide additional time to complete assignments. However, I must know something is going on to help. Please contact me as soon as you can; I don't need details – just a message that life has become difficult or complex and some idea of what this means for the course.

## Saving Your Work

If you are a current or potential statistics or data analytics major, please be sure to maintain both electronic and hard copies of your major assignments (e.g., projects) from this course. You may wish to refer to your class products during graduate school applications or job interviews. Additionally, for the senior comprehensive evaluation, one of the options for statistics majors is to put together a portfolio of their work including examples of past assignments. The work in this course could be a meaningful contribution to your statistics portfolio.

## Honor Code

Elon's honor pledge calls for a commitment to Elon's shared values of Honesty, Integrity, Responsibility, and Respect. To be clear about what constitutes violations of these values; students should review code of conduct policies described in the student handbook.

Students with questions about the specific interpretation of these values and violations as they relate to this course should contact this instructor immediately. Violations in academic-related areas will be documented in an incident report which will be maintained in the Office of Student Conduct and may result in a lowering of the course grade and/or failure of the course with an Honor Code F. Violations specifically covered by academic honor code policies include plagiarism, cheating, lying, stealing, and the facilitation of another's dishonesty.

## Policy on Recording Classes

Recorded classes may not be used in any way that denigrates and/or decontextualizes the instructor or any student whose class remarks are recorded. Any information contained in the recorded class may not be posted, published or quoted without the express consent of the instructor or speaker, and if permission is granted, must be properly cited. All recordings of class lectures or discussions are to be erased at the end of the semester.

## Intellectual Property Policy

With the incorporation of hybrid, blended and flipped classes, Elon faculty may choose to post written and/or recorded materials online that they typically would not post. It is important for all members of our community be familiar with Elon's Intellectual Property Policy.

## Disabilities Resources

Elon University is committed to providing an inclusive experience, accessible learning environments and equal opportunity to students with disabilities in accordance with the Americans with Disabilities Act and Section 504 of the Rehabilitation Act. Students needing reasonable accommodations should contact Disabilities Resources at 336.278.6568 or disabilities@elon.edu.

## Belk Library

The librarians in Belk Library help students find appropriate resources for course assignments. Students can schedule an appointment to meet with a librarian or chat with a librarian on the library website.

## Religious Holidays Policies

In supporting religious diversity, Elon has a policy and procedures for students who wish to observe religious holidays that are in conflict with the academic calendar, allowing students an excused absence. Students who wish to observe a holiday during the semester must complete the online Religious Observance Notification Form (RONF) by September 17, 2024. Students are required to make prior arrangements with the instructor for completion of any work missed during the absence. This policy does not apply during the final examination period. Once the completed RONF is received, the Truitt Center will send an e-mail to the instructor and the student that a RONF has been submitted. Students may contact the Truitt Center with questions.

## Writing Center

Elon's Writing Center in the Center for Writing Excellence is staffed by well-trained peer consultants who help with writing projects. The Writing Center is a welcoming place for all writers, regardless of experience or expertise. Writing Center consultants work with students on any kind of writing at any stage of the writing process. The Writing Center has two multimedia production studios to create and work with a consultant on multimedia and visual texts. Their presentation practice room allows students to record themselves giving a presentation and to get feedback from a consultant on accompanying visuals.

## Supporting Fellow Students in Distress

Occasionally, students may become worried or concerned about a fellow classmate's well-being. When this is the case, students are encouraged to share these concerns with their faculty and/or through Elon's Share a Concern website. Please see the Student Care and Outreach website for links to additional student resources such as identity-based resources, food and housing insecurity, getting around after an injury, and medical and personal leave

<u>Student Options Related to Enrollment in The Course</u>

**<u>Enrollment</u>**

Students should confirm their enrollment in courses through their OnTrack account. Students who do not appear on the course roll or do not show the correct course/section listed on OnTrack should consult with their instructor immediately.

**<u>Dropping a Course</u>**

Students may drop a course during the designated drop/add period through OnTrack. A course dropped during the drop/add period will not appear on the student's transcript.

**<u>Withdrawing from a Course</u>**

After the designated drop/add period, students may [withdraw from a course](#) without penalty before the course withdrawal deadline published in the Academic Calendar. Withdrawing from a course during this period will result in a mark of "W" that will appear on the student's academic transcript. Students may withdraw online via OnTrack. Students should not assume that they will be officially withdrawn from a course based on failure to attend class or notifying a faculty member of their intent to withdraw. The student is responsible for following the official process of withdrawing from a class.

<u>E-Warnings</u>

E-warnings are crucial for communicating with students and advisors when concerns arise about a student's academic performance. Advisors are copied on every e-warning submitted, which allows them to reinforce support. E-warnings are part of a standard process and not a personal judgement.

<u>Additional information on university course policies and dates</u>

Please consult the [Academic Catalog](#) and the [Academic Calendar](#).