# ST 558-601 Course Syllabus

## Data Science for Statisticians

### Fall 2025

Table of Contents

|  |
| --- |
| [ST 558-601 Course Syllabus](#_gjdgxs)  [INSTRUCTOR INFORMATION](#_t2zhy6vi46nt)  [Virtual Office Hours](#_5apq56vhpxvr)  [Preferred Method of Communication](#_160oipehe4tl)  [Response Time](#_9tq5evg7kj0r)  [COURSE INFORMATION](#_mqutz39h7baf)  [Meeting Time and Tool Used](#_7on7xe5icq8n)  [Prerequisites/Corequisites](#_tw03akvdflxh)  [COURSE OVERVIEW](#_cfn22wyg886q)  [Catalog Description](#_si9rkv8by90u)  [Structure](#_73c1r5flhbtz)  [LEARNING OUTCOMES](#_ejvqn9zczc0s)  [COURSE MATERIALS](#_hwu2947i7n80)  [Required Textbook and/or Software](#_u5qpjl5liw3a)  [Minimum Technical and Digital Information Literacy Skills](#_kmv6o1v81qm1)  [Hardware](#_4ku0hk2gulwu)  [Software](#_szgozn2yn2fm)  [Computer Skills and Digital Information Literacy](#_haumeihu7k50)  [GRADING](#_j0vvuywsdehw)  [Grading Policy](#_qcwc6ualllj8)  [Grading Scale](#_hj17funz8wq2)  [COURSE SCHEDULE](#_1416kxwpycrs)  [COURSE POLICIES](#_2bqnn9aste4t)  [Late Assignments](#_r3q8v9qmjsm0)  [Incomplete Grades](#_nra5401v6pja)  [Attendance and Participation](#_i5ql1bvfqmzd)  [General](#_hajpnrsvgjn8)  [UNIVERSITY POLICIES](#_hz67jackv27h)  [Academic Integrity and Honesty](#_tsalf89qn7nt)  [Originality Checking Software](#_63ieammmfk5a)  [Class recording statement:](#_heu81emfd4uu)  [Class privacy statement:](#_qrqj6grvdfd8)  [Students with Disabilities](#_7uc5epv9t8l1)  [Safe at NC State](#_nvmoriimu7vc)  [Supporting Fellow Students in Distress](#_ov9f6ubdmr5u)  [COURSE EVALUATIONS](#_4v9immn3nkkf)  [REFERENCES](#_24425rxc4xfb)  [SYLLABUS MODIFICATION STATEMENT](#_msswexpcw0ng) |

## **INSTRUCTOR INFORMATION**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Office Phone** | **Email** | **Office Location** |
| Justin Post | 919-515-0637 | jbpost2@ncsu.edu | SAS Hall 5272 |
| Ryan Li (Teaching Assistant) | NA | rli18@ncsu.edu | NA |

### Virtual Office Hours

My office hours are Mondays from 12:00-2:00pm and Thursdays from 1:00-2:00pm. The link is available on the course Moodle page.

* The Monday session is joint with another course so you may drop in while material from another class is being discussed - have no fear!
* The Thursday session is a ‘live’ session where we’ll answer questions and work through some problems. This is optional of course and will be recorded and posted for those that can’t make it.
* I’m also happy to hold office hours by appointment if those times don’t work for you (and the TA times don’t work). Just reach out whenever you’d like to meet. I’m pretty flexible!

The teaching assistant (TA), Ryan Li, will have zoom office hours on Tuesdays from 12:00-2:00pm ET.

### Preferred Method of Communication

The general discussion board is the preferred method of communication for asynchronous content questions and email for everything else.

### Response Time

You should receive a response within one business day at the latest (generally it should be shorter than that during business hours).

## **COURSE INFORMATION**

**Course Website**: <https://wolfware.ncsu.edu/courses/my-wolfware/> **Course Credit Hours**: 3

### Meeting Time and Tool Used

No in person meetings or synchronous activities (other than optional zoom office hours).

### Prerequisites/Corequisites

ST 555 (or equivalent), ST 511/ST 513/ST 517 (or equivalent)

## **COURSE OVERVIEW**

### Catalog Description

Methods for reading, manipulating, and combining data sources including databases. Custom functions, visualizations, and summaries. Common analyses done by data scientists. Methods for communicating results including dashboards. Regular access to a computer for homework and class exercises is required.

### Structure

The course is completely **asynchronous**, which means that students have no real-time class meeting requirements. Instead students will watch videos, complete and upload homework assignments/projects, and participate in forum posts. Students should set aside sufficient time in their schedules to complete these materials.

To obtain course help there are a number of options:

* Discussion Forum - This should be used for any question you feel comfortable asking and having others view. The TA, other students, and I can answer questions there. This will be the fastest way to receive a response!
* E-mail - If there is a question that you don't feel comfortable asking the whole class you can use e-mail. The TA and I will be checking daily (during the regular work week).
* Zoom Office Hour Sessions - These sessions can be used to share screens and have multiple users. You can do text chat, voice, and video. They are great for a class like this!

Each week will usually have videos to watch, readings, and an assignment to put what you’ve learned into practice. The notes and code used for each video are available as well.

## **LEARNING OUTCOMES**

**Course Learning Outcomes (COs):** At the end of this course students will be able to

* explain the steps and purpose of programs (CO 1)
* efficiently read in, combine, and manipulate data (CO 2)
* utilize help and other resources to customize programs (CO 3)
* write programs using good programming practices (CO 4)
* explore data and perform common analyses (CO 5)
* create reports, web pages, and dashboards to display and communicate results (CO 6)

## **COURSE MATERIALS**

### Required Textbook and/or Software

Textbooks: None - we'll use some free online texts and articles. Links to relevant readings appear in the weekly overviews.  
  
Software: Students in this course will use the [R statistical software](http://www.r-project.org/) and the [R Studio IDE](https://www.rstudio.com/). This software is open source, works on all major platforms, and is free to anyone. It is widely used in statistics and data science (behind python in data science but still quite popular).

Other than off campus proctoring of exams, there are no required costs for students in the course.

## **Minimum Technical and Digital Information Literacy Skills**

### Hardware

NC State’s Online and Distance Education provides [technology requirements and recommendations](https://online-distance.ncsu.edu/get-started/technology-requirements/) for computer hardware.

Access to a computer that you are able to install software on is required for the course.

### Software

* [Moodle and Wolfware](https://delta.ncsu.edu/knowledgebase/using-moodle-with-wolfware/)
  + [Moodle Accessibility Statement](https://docs.moodle.org/34/en/Accessibility)
  + [Moodle Privacy Policy](https://moodle.org/mod/page/view.php?id=8148)
  + [NCSU Privacy Policy](https://www.ncsu.edu/privacy/)
* [Adobe Reader](https://get.adobe.com/reader/) (for reading PDF files)
  + [Accessibility Statement](https://wwwimages2.adobe.com/content/dam/acom/en/accessibility/pdfs/accessing-pdf-sr.pdf)
  + [Adobe Privacy Policy](https://www.adobe.com/privacy/policy.html)
* [Zoom](https://support.zoom.us/hc/en-us):
  + [Zoom Accessibility Statement](https://zoom.us/accessibility)
  + [Zoom Privacy Policy](https://zoom.us/privacy)
* [G Suite](https://google.oit.ncsu.edu/)
  + [Accessibility Statement](https://support.google.com/a/answer/1631886?hl=en)
  + [Privacy Policy](https://edu.google.com/k-12-solutions/privacy-security/?modal_active=none)
* [Office 365](https://oit.ncsu.edu/my-it/hardware-software/software/office-365/)
  + [Accessibility Statement](https://www.microsoft.com/en-us/accessibility/office)
  + [Privacy Policy](https://products.office.com/en-us/business/office-365-trust-center-privacy)
* [R Statistical Software](https://cran.r-project.org/) and [R Studio](https://rstudio.com/) for programming
* [Miktex](https://miktex.org/) distribution for creation of PDF files
* [Docker](https://www.docker.com/) for creating containers
* [Github](http://github.com) for version control and collaboration. We recommend the use of the github rather than NCSU’s github.
* Headsets with microphone (optional for office hours)

### Computer Skills and Digital Information Literacy

The prerequisite coursework requires that students have completed a programming course. This implies that students are expected to have a strong understanding of how computers function and the logic required to instruct them.

Often in homework assignments students will be asked a few questions that require outside of class material. This implies that students are expected to be able to search through online help forums, vignettes, and the like to distill and extract relevant information to solve problems.

## **GRADING**

### Grading Policy

It is the student’s responsibility to be aware of their grades in the course and the appropriate level of work required. The **Moodle gradebook final grades are likely not correct** so be sure to keep track of how you are doing in the course. Your final grade depends on the following (subject to change with notification):

|  |  |
| --- | --- |
| **Item** | **Portion of Grade** |
| **Homework (8)** | **16% in total** |
| **Projects (3)** | **44% (14%, 14%, and 16% for the third project)** |
| **Exams (2)** | **40% (20% each)** |

**Homework:** There will be nine homework assignments (one dropped) that give you a low stakes chance to practice programming. There will also be some explanation questions included in these to help you practice for the exam. These assignments each have equal weight. **Parts of these assignments will not be graded for correctness.** Example keys will be posted for each assignment. **It is the student’s responsibility to make sure they understand what they did successfully and unsuccessfully on their homework assignments.**

**Projects:** There will be three larger projects with the last project acting as the final for the course. Some of these projects will require collaboration with others (data science is often a collaboration!). Details will follow as we get closer to the projects being assigned. Grading guides/rubrics will be provided with the assignments. No late work will be accepted for these - please contact me immediately if you have something unavoidable come up.

**Homework and projects can be turned in by 7:59am the day after they are due with no penalty. (The link stays open until then and you can submit.)**

**Exams:** All exams are **closed book and closed notes**. The exams require an in-person proctor.

Exam 1 window: T-Th, 9/23-9/25  
Exam 2 window: W-F, 11/19-11/21

Students who are unable to take an exam during the designated exam window for a legitimate unavoidable reason may be given the opportunity for a make-up exam or be given a reweighting of their scores. The exams are limited to 75 minutes. Many of the questions will be multiple choice, fill in the blank, or true/false. Other questions will be in essay form. These are generally graded on a 0/1 or 0/1/2/3 basis.

Please find information about taking exams with DELTA below:

**DELTA Testing Services** — This course requires proctored exams facilitated through [DELTA Testing Services](http://go.ncsu.edu/testing). A proctor is an impartial third party who verifies the identity of the student and ensures the academic integrity of an exam.

**Local students** —DELTA Testing Services will offer exams for on-campus students at one of the local test centers. Please visit the DELTA Testing Services website for [more information about on-campus testing](https://testing-services.delta.ncsu.edu/testing-services-on-campus/).

* [**Step 1: Make an Appointment**](https://testing-services.delta.ncsu.edu/testing-services-on-campus/on-campus-testing-guidelines/). Exams at the DELTA Test Centers are by **appointment only**.
* **Step 2: Come Prepared**.
  + Bring a physical photo ID
  + Know your Unity ID
  + DUO, bring your registered device
  + **Note pages will be collected after the exam. Make a copy of any notes before coming to the test center.**

**Students with Accommodations**— If you have approved accommodations with NC State’s Disability Resource Office (DRO), DELTA Testing Services wants to ensure that you receive the appropriate accommodations when you go to the test center.

* When scheduling at the test centers, you will see your approved accommodations listed as part of your appointment.
* Students who plan to use their accommodations at an off-campus location should provide the accommodation information to their proctors.

**Off-Campus students, including Meazure Learning/ ProctorU** — DELTA Testing Services, will oversee the process of approving an off-campus proctor, sending all exam materials, and receiving any materials from your proctor. Students are responsible for any additional fees associated with off-campus proctoring. To request off-campus proctoring, visit [Off-Campus Testing](https://testing-services.delta.ncsu.edu/testing-services-remote/).

* DELTA Testing Services certifies Meazure Learning/ ProctorU to administer exams for students taking fully digital exams. Paper-based exams are not permitted. Guidelines and procedures can be located on the [website](https://testing-services.delta.ncsu.edu/testing-services-remote/online-proctoring/).

If you have any additional questions, contact delta-testing@ncsu.edu or call 919.513.1513.

**Feedback Schedule:** The projects, and exams are graded manually. My goal as an instructor is to provide feedback and a grade within 5 business days of when you turn these assessments in. However, sometimes it might take a little longer (especially for projects).

### Grading Scale

This course uses this grading scale:

|  |  |  |
| --- | --- | --- |
| **Low** | **Letter** | **High** |
| 98 ≤ | A+ | ≤ 100 |
| 93 ≤ | A | < 97 |
| 90 ≤ | A- | < 93 |
| 88 ≤ | B+ | < 90 |
| 83 ≤ | B | < 87 |
| 80 ≤ | B- | < 83 |
| 78 ≤ | C+ | < 80 |
| 73 ≤ | C | < 77 |
| 70 ≤ | C- | < 73 |
| 68 ≤ | D+ | < 70 |
| 63 ≤ | D | < 67 |
| 60 ≤ | D- | < 63 |
| 0 ≤ | F | < 60 |

Note: This course may not be taken as an **S/U graded course**.

Students who wish to **audit** the course with satisfactory status must register officially for the course and will be required to complete 7 of the 9 homework assignments and complete at least 2 of the 3 projects with a 70% or better on average to receive credit.

## **COURSE SCHEDULE**

A full schedule is available on the [github site for the course](https://jbpost2.github.io/ST-558-Data-Science-for-Statisticians/).

Please note: course schedule is subject to change.

## **COURSE POLICIES**

### Late Assignments

Assignments can be turned in by 7am following the day they are due with no penalty. If you have a large conflict in which you are unable to complete an assignment or project prior to the due date, please contact Dr. Post immediately to work out other arrangements.

### Incomplete Grades

Incomplete (IN) grades are given only as specified in university regulations.

### Attendance and Participation

There are generally no attendance or participation requirements. Attendance is really done through your watching of videos and participation is done through your discussion board posts and assignments.   
NC State’s Attendance Policy:[*https://policies.ncsu.edu/regulation/reg-02-20-03-attendance-regulations/*](https://policies.ncsu.edu/regulation/reg-02-20-03-attendance-regulations/)Withdrawl Process:[*https://studentservices.ncsu.edu/your-classes/withdrawal/process/*](https://studentservices.ncsu.edu/your-classes/withdrawal/process/)

### General

Be nice! Also, some things in this course are going to be hard - technology doesn't always work. Be prepared to troubleshoot, check stackoverflow and other online help resources to get things working. **Many homework assignments will include some new aspects that will require you to problem solve (just like real life!).**

## **UNIVERSITY POLICIES**

### Academic Integrity and Honesty

Students are required to comply with the university policy on academic integrity found in the [Code of Student Conduct 11.35.01 sections 8 and 9](http://policies.ncsu.edu/policy/pol-11-35-01). Therefore, students are required to uphold the Pack Pledge: “I have neither given nor received unauthorized aid on this test or assignment.” Violations of academic integrity will be handled in accordance with the [Student Discipline Procedures](https://policies.ncsu.edu/regulation/reg-11-35-02/).

Please refer to the [Academic Integrity](https://studentconduct.dasa.ncsu.edu/academic-integrity-overview/) web page for a detailed explanation of the University’s policies on academic integrity and some of the common understandings related to those policies.

#### Originality Checking Software

Software may be used in this course to detect the originality of student submissions. Any work submitted is expected to be work completed solely by the student (not by large language models, chatbots, etc.)

#### Class recording statement:

Students may be able to be identified in the problem session/live session recordings. These videos may be shared with other classes at NC State in future semesters.

#### Class privacy statement:

This course requires online exchanges among students and the instructor, but NOT with persons outside the course. Students may be required to disclose personally identifiable information to other students in the course, via electronic tools like email or web postings, where relevant to the course. Examples include online discussions of class topics and posting of student coursework. All students are expected to respect the privacy of each other by not sharing or using such information outside the course.

Students are responsible for reviewing the NC State University PRR’s which pertains to their course rights and responsibilities:

* [Equal Opportunity and Non-Discrimination Policy Statement](https://policies.ncsu.edu/policy/pol-04-25-05) and [additional references](https://oied.ncsu.edu/equity/policies)
* [Code of Student Conduct](https://policies.ncsu.edu/policy/pol-11-35-01)
* [Grades and Grade Point Average](https://policies.ncsu.edu/regulation/reg-02-50-03)
* [Credit-Only Courses](https://policies.ncsu.edu/regulation/reg-02-20-15)
* [Audits](https://policies.ncsu.edu/regulation/reg-02-20-04)

**Students should refrain from using chatbots and generative AI to do their work. If you use these and simply copy and paste code, this is considered an academic integrity violation (and you likely won’t be able to handle the projects and the exams will be more difficult). You can use generative AI to help you debug, get ideas, etc. However, again, you cannot use them to create your code for you.**

**Similarly, sites like stackoverflow are great for debugging or figuring certain things out. It is ok to use them to help you debug, get ideas on fixes, etc. However, the code you turn in should be your code. You must understand it and be able to explain it.**

Any student found to submit code from one of these sources will receive a zero for the associated assignment and be directed to the academic integrity program here at NC State.

### Students with Disabilities

Reasonable accommodations will be made for students with verifiable disabilities. In order to take advantage of available accommodations, students must register with the [Disability Resource Office (DRO)](https://dro.dasa.ncsu.edu/). For more information on NC State’s policy on working with students with disabilities, please see the [Policies, Rules and Regulations page maintained by the DRO](https://dro.dasa.ncsu.edu/about-us/policies-rules-regulations/) and [REG 02.20.01 Academic Accommodations for Students with Disabilities](https://policies.ncsu.edu/regulation/reg-02-20-01/).

### Safe at NC State

At NC State, we take the health and safety of students, faculty and staff seriously. [The Office of Equal Opportunity](https://equalopportunity.ncsu.edu/) supports the university community by providing services and resources to support and guide individuals in obtaining the help they need. See the [Safe at NC State webpage](https://equalopportunity.ncsu.edu/safe/) for resources.

### Supporting Fellow Students in Distress

As members of the NC State Wolfpack community, we each share a personal responsibility to express concern for one another and to ensure that this classroom and the campus as a whole remain a healthy and safe environment for learning. Occasionally, you may come across a fellow classmate whose personal behavior concerns or worries you, either for the classmate’s well-being or yours. If you feel this way, I would encourage you to report this behavior to the [NC State CARES website](https://prevention.dasa.ncsu.edu/nc-state-cares/about/). Although you can report anonymously, it is preferred that you share your contact information so they can follow up with you personally.

## **COURSE EVALUATIONS**

ClassEval is the end-of-semester survey for students to evaluate the instruction of all university classes. The current survey is administered online and includes 12 closed-ended questions and 3 open-ended questions. Deans, department heads, and instructors may add a limited number of their own questions to these 15 common-core questions.

Each semester students’ responses are compiled into a ClassEval report for every instructor and class. Instructors use the evaluations to improve instruction and include them in their promotion and tenure dossiers, while department heads use them in annual reviews. The reports are included in instructors’ personnel files and are considered confidential.

Online class evaluations will be available for students to complete during the last two weeks of the semester for full-semester courses and the last week of shorter sessions. Students will receive an email directing them to a website to complete class evaluations. These become unavailable at 8 am on the first day of finals.

* Contact ClassEval Help Desk: [classeval@ncsu.edu](mailto:classeval@ncsu.edu)
* [ClassEval website](http://go.ncsu.edu/cesurvey)
* [More information about ClassEval](http://oirp.ncsu.edu/surveys/classeval)

## **REFERENCES**

Below is a list of references for the material in the course

Books:

* Wickham, H. and Grolemund, G. (2016). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O’Reilly Media
* Wickham, H. (2015). *Advanced R*. Chapman & Hall/CRC the R series
* R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing
* Wickham, H. (2015). *R packages*. O’Reilly Media
* Bryan, J., the STAT 545 TAs, Hester, J. (2020). *Happy Git and GitHub for the useR.*
* James, G., Witten, D., Hastie, T, and Tibshirani, R. (2017). *An Introduction to Statistical Learning.* Spring Texts in Statistics
* Hastie, T., Tibshirani, R., and Friedman, J. (2017). *The Elements of Statistical Learning.* Springer Series in Statistics

Websites:

* <https://medium.com/odscjournal/data-scientists-versus-statisticians-8ea146b7a47f>
* <https://www.springboard.com/blog/machine-learning-engineer-vs-data-scientist/>
* <https://www.simplilearn.com/data-science-vs-data-analytics-vs-machine-learning-article>
* <https://mixpanel.com/blog/2016/03/30/this-is-the-difference-between-statistics-and-data-science/>
* <https://www.smashingmagazine.com/2014/08/build-blog-jekyll-github-pages/>
* <https://guides.github.com/activities/hello-world/>
* <https://cran.r-project.org/>
* <https://support.rstudio.com/hc/en-us/articles/201057987-Quick-list-of-useful-R-packages>
* <https://www.r-bloggers.com/consistent-naming-conventions-in-r/>
* <https://google.github.io/styleguide/Rguide.html>
* <https://datacarpentry.org/R-ecology-lesson/05-r-and-databases.html>
* <https://db.rstudio.com/dplyr/>
* <https://www.w3schools.com/sql/>
* <https://api.census.gov/data.html>
* <https://cran.r-project.org/web/packages/censusapi/vignettes/getting-started.html>
* <https://www.programmableweb.com/apis/directory>
* <https://www.programmableweb.com/news/how-to-access-any-restful-api-using-r-language/how-to/2017/07/21>
* <https://gitforwindows.org/>
* <https://osxdaily.com/2014/02/12/install-command-line-tools-mac-os-x/>
* <http://www.sthda.com/english/wiki/descriptive-statistics-and-graphics#descriptive-statistics-for-a-single-group>
* <http://www.parallelr.com/r-with-parallel-computing/>
* <https://www.r-bloggers.com/scheduling-r-markdown-reports-via-email/>
* <https://tensorflow.rstudio.com/guide/tfestimators/estimator_basics/>
* <https://tensorflow.rstudio.com/guide/keras/>
* <https://www.asimovinstitute.org/neural-network-zoo/?utm_content=buffer09a5e&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer>
* <https://pathmind.com/wiki/neural-network>
* <https://medium.com/tmobile-tech/r-can-api-c184951a24a3>
* <https://medium.com/tmobile-tech/using-docker-to-deploy-an-r-plumber-api-863ccf91516d>
* <https://docker-curriculum.com/>
* <https://colinfay.me/docker-r-reproducibility/>

## **SYLLABUS MODIFICATION STATEMENT**

Some modifications may be needed to the course as we go along. If the syllabus is modified in a substantial way, students will be contacted via an announcement.