

ECON 5253: Data Science for Economists (Spring 2026)

Tyler Ransom

Email	ransom@ou.edu
Office	322 CCD1
Office Hours	Tu 12:30-1:30pm, W 2:00-3:00pm and by appointment
GitHub	tyleransom

- **Meeting day/time:** T,Th 1:30-2:45pm, CCD1, Room 338
- Office hours also available by appointment
- This course takes inspiration and extensively borrows materials from similar courses taught by [Jason DeBacker](#) (U of South Carolina), [Rick Evans](#) (Rice U), and [Grant McDermott](#) (U of Oregon). Thanks to them for providing a framework for using GitHub as a class collaboration tool and for insights into teaching programming skills.

Course description

Data science is a rapidly developing field that combines the recent Big Data revolution with ever-developing statistical algorithms to inform business and policy decisions. Nearly every company you've heard of uses data science to optimize its services: Netflix uses it to recommend new programs to its viewers, Amazon uses it to determine how much it should charge for its Prime services. This class will provide students with an overview of the data science workflow, from collecting raw data to drawing a set of insights from which a decision maker can make informed decisions. Along the way we will broadly cover a variety of advances in data collection, data storage, visualization, machine learning and econometrics topics, as well as teaching and reinforcing good programming practices. The primary goal of this course is to provide you, the student, with a set of skills that will allow you to compete for a data science job.

Course Objectives and Learning Outcomes

By the end of the course, students should be able to do the following:

1. Explain the data science workflow from start to finish
2. Be able to collect data from online sources via APIs or scraping
3. Describe similarities and differences between econometrics and machine learning
4. Explain what data science is, and how Big Data differs from other types of data
5. Demonstrate good programming practices by writing code that can allow for easy collaboration with others
6. Understand the differences between prediction and causality, and the cases in which each is useful

In this course students, through lecture and application, will learn about:

- * Good programming practices, including how to write code collaboratively with others
- * Software to increase research productivity including:

 - * LaTeX/Markdown
 - * git
 - * Software to collect & clean data, and estimate statistical models:

 - * R
 - * Julia
 - * Python

 - * Software to manage big data sets:

 - * SQL
 - * RDDs (Resilient Distributed Datasets) — Spark, Hadoop

 - * How to access and utilize cluster computing resources
 - * SSH (Secure Shell)
 - * SFTP (Secure File Transfer Protocol)
 - * SLURM (Simple Linux Utility for Resource Management)

- * Methods to gather and handle data including:

 - * Costs and benefits of different data structures
 - * Using APIs
 - * Web scraping
 - * Best practices for cleaning and visualizing data
 - * Computational methods to:

 - * Optimize and find roots of functions
 - * Perform Monte Carlo simulations
 - * Run computations in parallel using multiple processors (time permitting)

 - * Basics for modeling different types of data
 - * Machine learning basics:

 - * Supervised vs. unsupervised learning
 - * The five “tribes” of machine learning: how they are interconnected, and how they differ
 - * Machine learning vs. econometrics: prediction vs. causality
 - * Evaluating model performance
 - * Using economic models to inform policy decisions
 - * Computing structural models

Grades

Grades will be based on the categories listed below with the corresponding weights.

Component	Percent
Class Participation	10%
Problem Sets	35%
Exam & Quizzes	20%
Final project	35%
Total points	100%

Final grades will be assigned according to the standard cutoffs (90%+ for an A, 80%-89.99% for a B, etc.).

- **Participation:**
 - An important part of learning is face-to-face interaction. Thus, some of your grade will depend on attendance and active participation in class meetings.
- **Problem sets:** will be assigned approximately weekly throughout the semester.
 - You must write and submit your own computer code, although I encourage you to collaborate with your fellow students. I **DO NOT** want to see a bunch of copies of identical code. I **DO** want to see each of you learning how to code these problems so that you could do it on your own.
 - Problem set solutions, both written and code portions, will be turned in via a pull request from your private [GitHub.com](https://github.com) repository which is a fork of the class master repository on my account. (You will need to set up a GitHub account if you do not already have one.)

- Written solutions must be submitted as PDF documents or Jupyter Notebooks.
- Problem sets will be due on the day listed in the Daily Course Schedule section of this syllabus (see below) unless otherwise specified. Late problem sets will not receive any credit. Partially completed problem sets will receive partial credit.
- **Exam & Quizzes:**
 - We may periodically have in-class quizzes as low-stakes ways to get feedback
 - There will be an oral final exam, but no midterm exams
- **Final Project:**
 - Collect data on and analyze a research question of your choosing, using methods taught in this course
 - Write up a ~10 page (12pt font, double spaced, excluding References, Figures, and Tables) summary of your findings, including discussion about what prior studies of the same topic have found, as well as citations to prior studies
 - Turn in the written summary report and a GitHub repository containing all materials required to reproduce the results
 - Summary report should be written in LaTeX or RMarkdown and turned in as a PDF (source code for the summary report should also be included in your GitHub repository)
 - An example of what the final product should look like is [here](#), with LaTeX source code [here](#) and BibTeX source code [here](#).
 - A detailed rubric for the final project is [here](#)

Communication

- I will always be available via email, and via Zoom during office hours. Zoom link for office hours will be posted on Canvas.

Daily Course Schedule

(Will be continuously updated throughout the semester)

Date	Day	Topic	Due
Jan 20	T	What is data science / big data / why is it important? (Slides)	
Jan 22	Th	Git, GitHub, computing environment, and Coding best practices (Notes on computing enviornment); Slides on Git by Grant McDermott; Slides on clean coding by Tyler Ransom	Read Gentzkow & Shapiro's handbook ; Ch. 1 of <i>The Master Algorithm</i> ; register for GitHub account

Date	Day	Topic	Due
Jan 27	T	Linux command line (Grant McDermott's slides), SSH, accessing OSCER (Notes); Git Tutorial (p. 19 here ; adding upstream repositories here)	
Jan 29	Th	Overview of Data Scientists' tools (Slides)	PS 1
Feb 3	T	Using data: data types, storage (Slides)	
Feb 5	Th	Big Data: SQL (Slides) & RDDs (link)	
Feb 10	T	Sampling & storing Big Data (Slides); running jobs on the OSCER cluster	PS 2
Feb 12	Th	Web scraping/APIs to gather data (Grant McDermott's Lecture Notes; Ethics in Web Scraping ; rvest demonstration slides at 2018 useR conference; tidyverse cheat sheet; Grant McDermott's Lecture Notes on R language basics)	Sign up for API key at FRED and Data.gov
Feb 17	T	Web scraping/APIs to gather data (Grant McDermott's Lecture Notes)	PS 3
Feb 19	Th	Intro to Julia (Slides ; Julia's "Learning Julia" page)	
Feb 24	T	ggplot2 (Basics ; Kieran Healy's book)	PS 4
Feb 26	Th	Getting to know your data: descriptive statistics, cleaning, tips, tricks, transformations, visualization (Slides)	
Mar 3	T	Modeling continuous and discrete variables (Slides ; Simple R script)	
Mar 5	Th	Using JuMP to optimize cool stuff (Jupyter Notebook ; Julia Code) (in previous years: Linear Algebra Introduction / Review (Handout))	
Mar 10	T	Introduction to optimization (Notes)	PS 5
Mar 12	Th	No Class (Prof. Ransom travel)	
Mar 17	T	No Class (Spring Break)	
Mar 19	Th	No Class (Spring Break)	
Mar 24	T	Writing and optimizing functions in R, Python, and Julia (Slides)	PS 6
Mar 26	Th	Writing and optimizing functions in R, Python, and Julia (Slides)	
Mar 31	T	Debugging strategies and simulations (Slides)	PS 7

Date	Day	Topic	Due
Apr 2	Th	Intro to Machine Learning (Slides)	
Apr 7	T	Training and validating models with <code>tidymodels</code> (Slides)	PS 8
Apr 9	Th	Supervised ML: The 5 Tribes of Machine Learning (Slides)	
Apr 14	T	Unsupervised ML: Clustering (Slides)	PS 9
Apr 16	Th	Unsupervised ML: Dimensionality reduction and reinforcement learning (Slides)	
Apr 21	T	Machine learning vs. econometrics (Slides)	PS 10
Apr 23	Th	Discrete choice modeling: static and dynamic discrete choice (Slides)	
Apr 28	T	Strategies for more powerfully utilizing LLMs (Slides)	PS 11
Apr 30	Th	Final Project presentations (Rubric)	
May 5	T	Final Project presentations (Rubric)	
May 7	Th	Final Project presentations (Rubric)	PS 12
May 11	M		Final project due (Scoresheet)

Helpful Links

- [QuantEcon](#)
- [Notes on Machine Learning & Artificial Intelligence](#) by Chris Albon
- [R data wrangling cheatsheet](#)
- [R tidyverse](#)
- [Julia vs. Python for Data Science](#)
- [Machine Learning “Mind Map”](#)
- [JP Morgan massive overview of Big Data & Machine Learning](#)
- [Why it’s becoming increasingly more difficult to learn to program](#)

Books

- [The Master Algorithm](#) ([Amazon link](#))
- [Julia for Data Science](#) ([Amazon link](#))
- [R for Data Science](#) ([Free PDF](#))
- [Data Science at the Command Line](#) ([Free eBook](#))

AI policy

You are allowed and encouraged to use artificial intelligence (AI) resources in all aspects of this course. This includes, but is not limited to, using AI to help you write code, to help you debug code, to help you find answers to questions, and to help you find data. You are also allowed and encouraged to use AI to help you with your problem sets and final project. You are not allowed to use AI on in-class assessments such as quizzes or exams unless explicitly permitted by the instructor.

University Policies

Mental Health Support Services

Support is available for any student experiencing mental health issues that are impacting their academic success. Students can either been seen at the University Counseling Center (UCC) located on the second floor of Goddard Health Center or receive 24/7/365 crisis support from a licensed mental health provider through TimelyCare. To schedule an appointment or receive more information about mental health resources at OU please call the UCC at 405-325-2911 or visit [University Counseling Center](#). The UCC is located at 620 Elm Ave., Room 201, Norman, OK 73019.

Title IX Resources and Reporting Requirement

The University of Oklahoma faculty are committed to creating a safe learning environment for all members of our community, free from sex-based discrimination, including sexual harassment, domestic and dating violence, sexual assault, and stalking, in accordance with Title IX. There are resources available to those impacted, including: speaking with someone confidentially about your options, medical attention, counseling, reporting, academic support, and safety plans. If you have (or someone you know has) experienced any form of sex-based discrimination or violence and wish to speak with someone confidentially, please contact OU Advocates (available 24/7 at 405-615-0013) or University Counseling Center (M-F 8 a.m. to 5 p.m. at 405-325-2911).

Because the University of Oklahoma is committed to the safety of you and other students, and because of our Title IX obligations, I, as well as other faculty, Graduate Assistants, and Teaching Assistants, are mandatory reporters. This means that we are obligated to report sex-based violence that has been disclosed to us to the Institutional Equity Office. This includes disclosures that occur in: class discussion, writing assignments, discussion boards, emails and during Student/Office Hours. You may also choose to report directly to the Institutional Equity Office. After a report is filed, the Title IX Coordinator will reach out to provide resources, support, and information and the reported information will remain private. For more information regarding the University's Title IX Grievance procedures, reporting, or support measures, please visit [Institutional Equity Office](#) at 405-325-3546.

Reasonable Accommodation Policy

The University of Oklahoma (OU) is committed to the goal of achieving equal educational opportunity and full educational participation for students with disabilities. If you have already established reasonable accommodations with the Accessibility and Disability Resource Center (ADRC), please log into iAdvise to request your semester accommodations as soon as possible and contact me privately, so that we have adequate time to arrange your approved academic accommodations.

If you have not yet established services through ADRC, but have a documented disability and require accommodations, please complete ADRC's pre-registration form to begin the registration process. ADRC facilitates the interactive process that establishes reasonable accommodations for students at OU. For more information on ADRC registration procedures, please review their website. You may also contact them at (405) 325-3852 or adrc@ou.edu, or visit www.ou.edu/adrc for more information.

Note: disabilities may include, but are not limited to, mental health, chronic health, physical, vision, hearing, learning and attention disabilities, pregnancy-related. ADRC can also support students experiencing temporary medical conditions.

Religious Observance

It is the policy of the University to excuse the absences of students that result from religious observances and to reschedule examinations and additional required classwork that may fall on religious holidays, without penalty.

Adjustments for Pregnancy and Related Issues

Should you need modifications or adjustments to your course requirements because of pregnancy or a pregnancy-related condition, please request modifications via the [Institutional Equity Office website](#) or call the Institutional Equity Office at 405-325-3546 as soon as possible. Also, see the [Institutional Equity Office FAQ on Pregnant and Parenting Students' Rights](#) for answers to commonly asked questions.

Final Exam Preparation Period

Pre-finals week will be defined as the seven calendar days before the first day of finals. Faculty may cover new course material throughout this week. For specific provisions of the policy please refer to OU's [Final Exam Preparation Period policy](#).

Emergency Protocol

During an emergency, there are official university procedures that will maximize your safety.

Severe Weather: If you receive an OU Alert to seek refuge or hear a tornado siren that signals severe weather:

1. Look for severe weather refuge location maps located inside most OU buildings near the entrances.
2. Seek refuge inside a building. Do not leave one building to seek shelter in another building that you deem safer. If outside, get into the nearest building.
3. Go to the building's severe weather refuge location. If you do not know where that is, go to the lowest level possible and seek refuge in an innermost room. Avoid outside doors and windows.
4. Get in, Get Down, Cover Up
5. Wait for official notice to resume normal activities.

Additional Weather Safety Information is available through the [Department of Campus Safety](#).

The University of Oklahoma Active Threat Guidance

The University of Oklahoma embraces a Run, Hide, Fight strategy for active threats on campus. This strategy is well known, widely accepted, and proven to save lives. To receive emergency campus alerts, be sure to update your contact information and preferences in the account settings section at one.ou.edu.

RUN: Running away from the threat is usually the best option. If it is safe to run, run as far away from the threat as possible. Call 911 when you are in a safe location and let them know from which OU campus you're calling from and location of active threat.

HIDE: If running is not practical, the next best option is to hide. Lock and barricade all doors; turn off all lights; turn down your phone's volume; search for improvised weapons; hide behind solid objects and walls; and hide yourself completely and stay quiet. Remain in place until law enforcement arrives. Be patient and remain hidden.

FIGHT: If you are unable to run or hide, the last best option is to fight. Have one or more improvised weapons with you and be prepared to attack. Attack them when they are least expecting it and hit them where it hurts most: the face (specifically eyes, nose, and ears), the throat, the diaphragm (solar plexus), and the groin.

Please save OUPD's contact information in your phone: * **NORMAN campus:** For non-emergencies call (405) 325-1717. For emergencies call (405) 325-1911 or dial 911. * **TULSA campus:** For non-emergencies call (918) 660-3900. For emergencies call (918) 660-3333 or dial 911.

Fire Alarm/General Emergency

If you receive an OU Alert that there is danger inside or near the building, or the fire alarm inside the building activates:

1. LEAVE the building. Do not use the elevators.
2. KNOW at least two building exits
3. ASSIST those that may need help

4. PROCEED to the emergency assembly area
5. ONCE safely outside, NOTIFY first responders of anyone that may still be inside building due to mobility issues.
6. WAIT for official notice before attempting to re-enter the building.

OU Fire Safety on Campus

Academic Integrity

I do not tolerate academic misconduct, [and neither does the University of Oklahoma](#). I will not hesitate to fail students who do not fully comply with the University's academic misconduct policy. If you find yourself contemplating cheating, plagiarism, or other forms of academic misconduct, please come see me first. Help is available if you are struggling. I want everyone in the class to try their best and to do their own work. Please be advised that I reserve the right to utilize anti-plagiarism resources such as TurnItIn when grading assignments.