

BUSN 5000

Introduction to Data Science for Business and Economics

Ian Schmutte

Chris Cornwell

Spring 2023
(updated 23 Dec 22)

Class Room: Amos B010

Class Hours: TR, 935a and 1110a

Teaching Team

Instructors

Contact	<i>Ian Schmutte</i>	<i>Chris Cornwell</i>
email	schmutte@uga.edu	cornwl@uga.edu
office	Amos B420	Amos B463
hours	by appt	W, 200-300

Teaching Assistants

Contact	<i>Morgan Kearns</i>	<i>Eric Gluckman</i>
email	morgan.kearns@uga.edu	eric.gluckman@uga.edu
office	Amos B415	Ivester E111

Coordinating GRA

All administrative questions about the course should be directed to the coordinating GRA, *Abbi Cormier*: abigail.cormier@uga.edu.

Course Description

The modern world is awash in a seemingly unlimited amount of data. To harness these data for decisions starts with acquiring the raw information and ends with a report describing the outcome of some analysis. At each step, the analyst combines data with some ideas about how the world works to produce an output. BUSN 5000 will take a hands-on approach, with a focus on techniques of data preparation; descriptive, explanatory and predictive analyses; and scientific communication.

Course Objectives

After completing this course, you should understand

1. how social, economic, and business data are produced from different sorts of information.
2. how to produce replicable, properly curated research results based on confidential and public-use data files.
3. how to map economic models, business problems, and policy questions to hypotheses about relationships in data.
4. the basic concepts of causal inference and how to implement basic research designs to facilitate causal inference.
5. the basic concepts of statistical learning and how to implement them to make predictions about economic, business or policy outcomes.
6. how to obtain data and prepare it for analysis.
7. how to describe data and perform basic descriptive, explanatory and predictive analyses.
8. how to communicate the results from descriptive, explanatory and predictive analyses.

Recommended Texts

There are no required texts for this course, but many useful ones. Here is a curated list where you can find the course content covered at an “age-appropriate” level:

- *Beginner*

Bueno de Mesquita, E. and Fowler, A. *Thinking Clearly with Data*, Princeton University Press.

Çetinkaya-Rundel, R., [Data Science in Box](#).

Healy, K., [Data Visualization: A Practical Introduction](#), Princeton University Press.

- *Next Level*

Angrist, J. and Pischke, S., *Mastering ‘Metrics*, Princeton University Press.

Cunningham, S., [Causal Inference: The Mixtape](#), Yale University Press.

Schwabish, J., *Better Data Visualizations*, Columbia University Press.

There are many excellent purely online resources to get a statistics refresher with some R programming tossed in. At the beginner level, we recommend [Computational and Inferential Thinking: The Foundations of Data Science](#), [R for Data Science](#) and [Learning Statistics with R](#). You will find a more advanced treatment in [Foundations of Statistics with R](#).

Software

The software of choice for this class is [R](#), a free and open-source language for statistical computing and graphics. [RStudio](#) is a popular integrated development environment (IDE) for R that will greatly enhance your R experience. First, [download](#) and install R; then [download](#) and install RStudio. Follow these [instructions](#).

Terry Analytics Lab

The TAL is a resource supported by the Department of Economics for all Terry students enrolled in business analytics and econometrics courses. The lab is managed by *Matt Wilson* (William.Wilson2@uga.edu). Her team provides free tutoring in course concepts and R coding. TAL typically operates from early/mid afternoon to early evening, Mon-Thu, beginning the second week of the semester. The TAs for this course are experienced TAL staff members with a dedicated assignment to BUSN 5000 students. We strongly encourage you to seek their assistance as you work on course assignments. TAL has its own eLC course page to which you will be subscribed.

Course Policies

Electronic devices

Cell phones *must* be muted or turned off and stowed away during class. Laptops may be used in class, but only for purposes directly related to the course (e.g., taking notes, live coding and viewing course materials).

Communication

Our communications to the class will generally come through the eLC Announcements tool, which functions like an instant messaging system. You should set your notifications preferences to receive Announcements postings in the manner that suits you. We strongly encourage the SMS option.

All questions about homework or the course material should first be directed to the course TAs. All questions about the course that are not about help with homework or understanding the course materials should first be directed to the coordinating GRA. If they are unable to solve your problem, they will advise you on where to go next.

Assistance with homework

For questions about course assignments, the TAs will hold regular office hours throughout the week via Zoom. You should consult them early and often to make sure assignments are completed on time.

Instructors

Prof. Cornwell has open-door drop-in hours on Wednesday from 2-3pm. Prof. Schmutte has office hours by appointment. You can schedule a meeting using Calendly <https://calendly.com/schmutte/office-hours>.

Performance evaluation

Your performance will be evaluated on the basis of homework assignments, Dailies, projects and a final exam weighted as follows:

Part	Assessment	Number	Weight	Total
I	Homework	4/5	.05	.20
	Dailies			.05
	Project	1	.15	.15
	Test	1	.10	.10
II	Homework	4/5	.05	.20
	Dailies			.05
	Project	1	.15	.15
	Test	1	.10	.10

Homework assignments are *formative* graded tutorials that guide you through the key concepts in each course topic and include an empirical component. We will count only the best 4 homework assignment scores from each part of the course. Homework assignment due dates are included in the schedules linked below. *Late homework submissions will not be graded. If you submit a homework assignment after the deadline, you will receive a grade of 0 for that assignment.*

Dailies are short active-learning exercises that we will conduct each class period. You will earn Dailies points by participating and providing correct responses. Your overall Dailies score will be the percentage of total Dailies points you earned. *If you are absent from class you cannot participate in the Daily.* (This probably goes without saying, but just in case: participating in Dailies on behalf of students who are absent from class constitutes an academic honesty violation and will be prosecuted as such.)

The projects are *summative* assignments covering the material in that part and include an extended empirical exercise in a notebook-style deliverable. In preparation for the projects, we may also provide occasional supplementary programming exercises. These are *not graded*. Their purpose is solely to help you build coding skills and develop code relevant for the projects.

The tests are *summative* assessments of the key concepts covering the material in each section of the course.

You will be ranked relative to other students in the class according to your overall performance and grades assigned based on your class rank. We will use the plus/minus system to make distinctions within grade categories.

Attendance

Regular class attendance is essential for success and therefore strongly encouraged. There is no explicit penalty assessed for missing class, but missing class means missing a Daily, which could be costly. In any event, repeated absences will send a clear negative signal. All instruction will be fully in-person. Class lectures and discussion will not be recorded.

Topical Outline

The topical outlines for parts I and II of the course are provided below. Follow the links to class schedules with reading assignments.

Part I [[class schedule with reading assignments](#)]

1. Reproducible data analysis
2. Documenting data and review of basic statistics
3. The Bayesian approach to learning from data
4. Linking theory and measurement
5. Sources of data error
6. Data ethics

Part II [[class schedule with reading assignments](#)]

1. Regression fundamentals
2. Potential outcomes and causal inference
3. Regression discontinuity
4. Difference in differences
5. Prediction with regression
6. Introduction to machine learning

Other Course Policies and Statements

UGA student Honor Code

“I will be academically honest in all of my academic work and will not tolerate academic dishonesty of others.” A Culture of Honesty, the University’s policy and procedures for handling cases of suspected dishonesty, can be found at <https://www.uga.edu/ovpi>. Every course syllabus should include the instructor’s expectations related to academic honesty.

Mental health and wellness resources

If you or someone you know needs assistance, you are encouraged to contact Student Care and Outreach in the Division of Student Affairs at 706-542-7774 or visit <https://sco.uga.edu>. They will help you navigate any difficult circumstances you may be facing by connecting you with the appropriate resources or services.

UGA has several resources for a student seeking mental health services (<https://www.uhs.uga.edu/bewelluga/bewelluga>) or crisis support (<https://www.uhs.uga.edu/info/emergencies>).

If you need help managing stress anxiety, relationships, etc., please visit BeWellUGA (<https://www.uhs.uga.edu/bewelluga/bewelluga>) for a list of FREE workshops, classes, mentoring, and health coaching led by licensed clinicians and health educators in the University Health Center.

Additional resources can be accessed through the UGA App.

Diversity, equity and inclusion

The Terry College is committed to promoting diversity, equity, inclusion, and belonging among its students, faculty, and staff. This class welcomes the open exchange of ideas and values freedom of thought and expression. This class provides a professional environment that recognizes the inherent worth of every person. It aims to foster dignity, understanding, and mutual respect among all individuals in the class.

Coronavirus information for students

For the latest on UGA policy, you can visit <https://coronavirus.uga.edu/>.

Changes to the syllabus

The course syllabus is a general plan for the course; deviations announced to the class by the instructor may be necessary.