

Teaching Complexity at Scale: Leveraging Technology to Get Back to Teaching

Abigail Cormier

John Munro Godfrey Sr. Department of Economics
University of Georgia

JETSET 2024
July 31, 2024



UNIVERSITY OF
GEORGIA

Starting out...

Take a moment and reflect on one challenging administrative/grading/burdensome task that keeps you away from time engaging with students.

What I am here to talk about today

How we asked the essential questions of the course and then I found and leveraged technological tools to:

- take on more students
- with a smaller staff
- increase high-impact in-person offerings to students
- increase active learning in the classroom
- increase equity through cloud computing

Importantly, it is about why these tools!

It was never about the camera

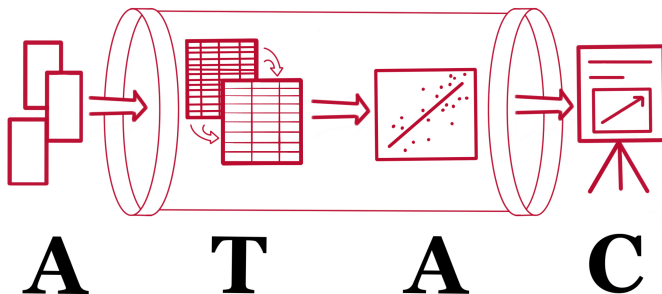


What is “Introduction to Data Science for Business and Economics”?

- Higher-level data-driven course
- We have to treat students as a blank slate and meet them where they are
- We help them get to communicating about causal inference and prediction

What do I mean by “at scale”:

- Spring '22: team of 9 teaching 70 → Fall '24: team of 2/2.5 teaching 250



How we got there:

It all started with 3 questions:

- What can we eliminate?
- What can we automate?
- What can we **not** leave behind?

What can we do now to increase engagement?

Thinking back to your problem, reflect on what technology would need to be able to do to decrease/eliminate the friction.

What it accomplished for students

- Instant feedback on their homework
- Near instant feedback (24-hr) on open-ended active learning quizzes
- Office hours where we work on the empirical homework together
- Office hours for project help
- Team-teaching in the classroom
- Live workshop of the communication skills of data science in-class
- Weekly recitation

Shinyapps for homework

Moving away from individual homework assignments coded on student's computers:

- High startup costs
- High amounts of frustration & confusion early in the course
- Huge grading burden
- Slow turn-around times for students (a week or more)

Into custom online assignments that

- Take the computing power into the cloud (equity)
- Check understanding and provide instant feedback
- Allow us to walk students through key coding skills

Shinyapps

The filtering action will take the original data set, `cps09mar`, filter on the age restriction, and create a new data set containing only the younger workers. The syntax goes like this:

```
newdataset <- olddataset %>%  
  filter(age <= _____,  
         age >= _____)
```

As an example, if we want filter out people younger than 25 (i.e., keep people who at least 25), the `filter` command should look like this:

```
newdataset <- olddataset %>%  
  filter(age >= 25)
```

Note that we have used `%>%`, or the “pipe”, to pass the data to the `filter` function. The pipe is very useful because it allows you to code a sequence of operations that is easy to read and interpret.

Now, use this sample code to complete the filtering operation to create a new data set called `cps09mar_2534` containing only the younger workers.

R Code [Start Over](#)

[Run Code](#)

[Submit Answer](#)

```
1 cps09mar_2534 <- _____  
2 _____(age <= _____,  
3           age >= _____)
```

We can also do open-ended coding where we check understanding after:

Now, plot the estimated CEFs just like in Figure 11, except the vertical axis should show actual dollar values.

R Code [↻ Start Over](#)

▶ Run Code

```
1 options(scipen=999) #get rid of scientific notation
2 ggplot(cef_fvm, aes(_____, _____, color=_____)) +
3   geom_point() +
4   geom_line() +
5   labs(title="CEFs of _____ by gender",
6         x = "Age",
7         y = "Average _____ by age")
```

Question 6:

The CEF plots indicate that the gender earnings gap (grows/shrinks) _____ over a typical career.

answer 1

Submit Answer

Manual grading, hard to quantify
performance

FIRST NAME AND LAST NAME: _____

SL# (DO NOT OMIT the "SL"): _____ DATE (dd/mm/yy): _____

BUSN 5000 :: Daily Form

You may do scratch work above the final answer boxes for each question.

Question 1

Q1 Final Answer:

Question 2

Q2 Final Answer:

Question 3

Q3 Final Answer:

Gradescope grading, easy to quantify

Gradescope

Within 24-hours, grades are posted along with solutions and performance summaries to return to at the start of class

Feb 8

One property of a good estimator is that it approaches the _____ as sample size increases.

estimand

Distribution of responses, $N = 190$

Response	Percent
estimand (correct)	93
population mean	1
sampling error	1
zero	1
estimator	1
not there	3

Gradescope

Within 24-hours, grades are posted along with solutions and performance summaries to return to at the start of class

Apr 2

Write a sentence that provides a clear interpretation of the **Female** coefficient estimate in the “Baseline” column of Table 6 of the project.

The baseline regression, which includes age and its square, produces a **Female** coefficient estimate of -0.162 , suggesting that women earn about 16.2 percent less than men on average.

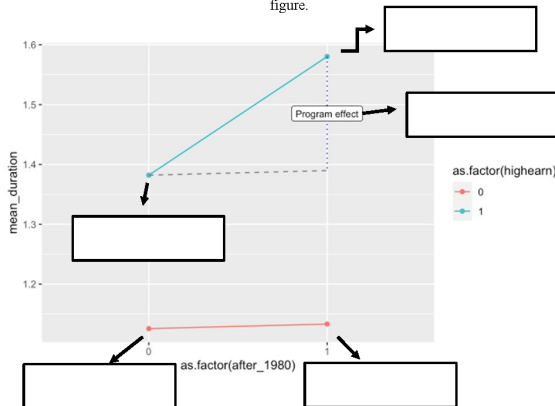
The baseline regression suggests that women earn about 16.2 percent less than men on average, controlling for age and its square.

Distribution of responses, $N = 200$

Response	Percent
Made the 16.2% point	74
Left it at -0.162	16
Not a sentence, wrong interpretation, no estimate included	11

Gradescope for exams beyond multiple choice

39. The figure below depicts the basic facts from the analysis of KY's worker's compensation benefit increase, namely the sample average log injury durations for high and low earners in KY before and after the increase. Use the estimands table in the exam booklet to label the five points on the figure.



Generative AI as a better than Gradescope option

	A	B	C
1	Number	Answer	GPT Feedback (First throw)
2		It is the average change (increase or decrease) in wages between men and women based on the two different variables: if the person is married, and if the person has children under 6.	Not quite: The response doesn't specifically report the findings from the correct model (Only Singles) and lacks clarity on the estimated gender wage gap for singles without children under 6, which was the focus of the prompt.
3		The baseline regression accounts for age and age-squared, which produces a female coefficient estimate of -0.162 , suggesting that females earn about 16.2% less than males on average.	Not quite: The response describes the baseline model instead of the model restricted to unmarried workers without children under 6. It fails to address the specific prompt regarding unmarried workers without children under 6.
4		By adding female and restricting the analysis to unmarried workers without children under 6 shows that on average, unmarried female workers without young children earn approximately 12.3% less than their male counterparts.	Good: This response is clear and accurate. It correctly identifies the wage gap for unmarried female workers without young children as approximately 12.3% less than their male counterparts, aligning well with the 'roughly right' answer.
5		Holding other factors constant, the presence of the Female variable shows a negative correlation between the gender and wage. Therefore, it is noted within the data and the regression that females earn less wages than males, and there is a wage gap between the two factors.	Not quite: This response provides a general statement about the relationship between gender and wages, but does not specify the wage gap for the model focused on unmarried workers without children under 6.

- Does not ever replace a human grader
- Does the same category first-throw that Gradescope does
- Except it can do it for open-ended questions!

Generative AI as teaching assistant

- CSS code help for nice visuals
- R code to help me automate the gradebook
- Python code help for video creation
- Assistance with evaluating homework question I want to revise
- It takes the tasks I can easily check, but do not want to spend energy doing

Generative AI as video creation assistant



**UNIVERSITY OF
GEORGIA**

Your turn!

Return to your one challenging task that keeps you away from time engaging with students:

- How could you leverage generative AI, an AI tool like Gradescope, or any other technology to alleviate or eliminate that task?

Now, let's start thinking creatively:

- What is one thing you have wanted to offer your students, but do not have the time to learn how to do it yourself?
- How could you use Generative AI or a different technology to help you accomplish this goal?

For what it's worth...

- 1 pt increase in “daily” average assoc. w/ 0.25 of an exam point
- Student success on our first coding project used to be an 80% submission rate
- Last semester, we were above a 97% submission rate!
- 12 percentage point increase in ability to correctly calculate a t -statistic
- 25 percentage point increase in ability to “plug-in” sample estimators for population objects

Anecdotes:

- “It is great. The TAs are really awesome and I feel I am learning a lot. It is a difficult class because it is so different from anything we’ve done before, but it is fun and useful.”
- “BUSN 5000 was really well organized which made the challenge manageable”