

# **ECON 337: Economics in the Age of Big Data**

## **University of Massachusetts Amherst**

### **Fall 2023**

Instructor: Ray Caraher

Last Updated: 10/08/2023

Email: [rcaraher@umass.edu](mailto:rcaraher@umass.edu)

Office Hours: Wednesday 2-4pm

Office: Gordon Hall room 106

Course page: <https://umamherst.instructure.com/courses/5398>

Class Hours: Tuesday/Thursday 4-5:15pm

Class Room: Machmer Hall room W-11

---

## **Course Description**

We are living in the age of Big Data.

In these "data-driven" times, the potential for addressing critical socio-economic issues related to education, health, poverty, climate change, and more is immense. However, it also brings to light serious ethical dilemmas and challenges. Throughout history, data has been a tool for social progress, but also as a means to justify social inequalities and the policies which exacerbate them. The era of Big Data is no different.

This course focuses on the practical applications of Big Data for social scientists and policymakers. We will learn the basic theory and skills for basic analysis of Big Data, and apply them to several of the pressing social problems facing the world today. However, we will also develop a critical perspective on the role of data in society and policy by focusing on historical and recent case studies of how data was used to justify social and economic inequalities. The course is divided into three modules. The first module is designed to introduce you to some of the historical misuses of data along with the political and ethical complexities associated with the rise of Big Data, Machine Learning, and AI. In the second module, you will gain practical knowledge and hands-on experience with data manipulation and analysis using the R programming language. In the last module, we will explore how Big Data has revolutionized the social sciences. We will delve into recent advancements that have expanded the possibilities of empirical analysis, enabling us to extract valuable insights from vast datasets.

## **Required Materials**

You are not required to purchase books for this course. All readings will be posted on Canvas. However, you must have a laptop or another way to access a computer to use R. Please contact me if you think this will be an issue.

## Prerequisites

ECON 103 (or RES-ECON 102) and ECON 104.

## Course Structure

This class will be a combination of lecture, discussion, and lab. We will also do several assignments in groups, so much of this work will be done with your classmates.

## Assessments

### Reading Notes

This class will be reading-intensive. It is required that you do all your readings the week that we cover the topic. We will primarily be reading book chapters and articles from academic journals. Starting the second week of class, you are required to take reading notes which highlight what you think is important in the text. These notes should be about one page for each reading using the format you think is best, and can be taken either electronically or handwritten. I will grade them on a three point scale. You will be required to submit these notes to Canvas at the end of every week by Thursday at 11:59pm. Reading notes must be submitted on time, but you will be allowed to miss three weeks of notes with no penalty. There will be no credit given for late reading notes in any circumstance.

### Discussion Questions

Starting the second week of class, 5 students per week are required to submit one discussion question based on the readings each prior to each class that week, which we will use to help guide our in-class discussions. You will be required to do this for 2-3 weeks during the semester. The discussion question will be due before class by 12:00pm on Tuesdays and Thursdays on Canvas. The discussion questions must reflect on the readings in a meaningful and critical way, and go beyond surface-level observations. You should also be prepared to elaborate on your discussion question during in-class discussions. I will grade them on a three point scale.

### Assignments

There will be four assignments in this class. One will primarily be short answer, two will be applying practical skills to real-life data, and the fourth will be focused on your group project. Some assignments will be done individually and some will be done in groups of 3-4 students. The format of each assignment will be discussed when it is assigned.

### Final Group Project

The final project for this class will be to research an interesting social problem related to the course material and present your findings during the last two weeks of class. You will pick from a list of topics I have prepared, but you may also suggest your own. These projects will be done in groups of 3-4. You will also be required to turn in a written report of your findings. Additional details about the final project will be discussed later in the semester.

## Participation

We will spend a substantial amount of time in this class discussing the assigned readings. You therefore must come to class having done the readings, prepared to discuss the course material, and ready to think through about the discussion questions in small groups or with the entire class. Rather than evaluating you based the number of times you contribute, your grade will instead focus on how your ability to critically engage with the readings in discussions grows throughout the semester.

## Grading

The standard UMass Amherst grading scale will be used. I reserve the right to curve the scale dependent on overall class scores at the end of the semester. Any curve will only ever make it easier to obtain a certain letter grade. The grade will count the assessments using the following proportions:

- 40% of your grade will be determined by the 4 assignments (10% each).
- 30% will be determined by your reading notes.
- 15% will be determined by your final project (10% for the presentation and 5% for the report).
- 10% will be determined by your discussion questions.
- 5% will be determined by your participation.

## Course Policies

### Attendance Policy

Unless otherwise noted, all classes will be in-person and attendance is required. You are allowed to miss up to three classes with no penalty if you let me know in advance. Any more than three absences will require a university-sanctioned excuse, or I will take one letter grade off your participation grade per missed class.

### Assignment Submission

Please submit all your assignments as either Microsoft Word files (.doc/.docx) or PDF files. All files must be submitted via Canvas (do not email them to me). You must confirm that your files can be opened/are not corrupt before you turn them in.

For reading notes, if you have been taking notes electronically submit them as Word or PDF files. If you have been taking hand-written notes, use your phone to take a picture or scan them. Please make sure they are legible before you turn them in.

### Late Assignments

All assignments must be completed on time. If for any reason you feel that you need a bit more time on an assignment, email me as soon as you make this determination and we can discuss possibly extending the deadline. Except in exceptional circumstances, extensions will not be granted if you notify me the same day. All unexcused late assignments will drop one letter grade per day. There will be no extensions granted in any circumstance for reading notes.

## **Academic Integrity and Honesty**

Since the integrity of the academic enterprise of any institution of higher education requires honesty in scholarship and research, academic honesty is required of all students at the University of Massachusetts Amherst. Academic dishonesty is prohibited in all programs of the University. Academic dishonesty includes but is not limited to: cheating, fabrication, plagiarism, and facilitating dishonesty. Appropriate sanctions may be imposed on any student who has committed an act of academic dishonesty. Instructors should take reasonable steps to address academic misconduct. Any person who has reason to believe that a student has committed academic dishonesty should bring such information to the attention of the appropriate course instructor as soon as possible. Instances of academic dishonesty not related to a specific course should be brought to the attention of the appropriate department Head or Chair. Since students are expected to be familiar with this policy and the commonly accepted standards of academic integrity, ignorance of such standards is not normally sufficient evidence of lack of intent. See <https://www.umass.edu/honesty/> for details.

## **Accommodations for Disabilities**

The University of Massachusetts Amherst is committed to providing an equal educational opportunity for all students. If you have a documented physical, psychological, or learning disability on file with Disability Services (DS), you may be eligible for reasonable academic accommodations to help you succeed in this course. If you have a documented disability that requires an accommodation, please notify me within the first two weeks of the semester so that we may make appropriate arrangements. For further information, please visit Disability Services at <https://www.umass.edu/disability/>. Even if you don't have a documented disability on file, you may still be eligible for accommodations. Please reach out to me at any time if you're struggling with some aspect of the course. There may be something I can do to help, or I may be able to direct you to the appropriate resources.

## **In-Class Recordings by Students and Unauthorized Distribution of Notes**

Students may only use the notes they take from class for their own personal use, or share with UMass Disabilities Services. Students cannot share or sell these notes via an outside vendor or entity without the faculty/instructor's permission. This pertains to in-class recordings as well. Usage of the notes or in-class recordings in this way without instructor permission is a violation of instructor copyright protection.

## Course Schedule

The schedule is tentative and subject to change. Readings we will discuss during the week are listed below.

**Week 1, 09/04 - 09/08:** Course Overview and Introduction

- D'Ignazio and Klein (2020), Introduction

### Module 1: Inequality, Power, and Science - Critical Perspectives on Big Data

**Week 2, 09/11 - 09/15:** What is Big Data and Is It Good or Bad?

- D'Ignazio and Klein (2020), Ch. 1, Ch. 6
- O'Neil (2016), Introduction, Ch. 1, Ch. 5

**Week 3, 09/18 - 09/22:** Historical Misuses of Data: Scientific Racism

- Gould (1996), Ch. 1, Ch. 5
- **Assignment 1 due Friday, Sept. 22 at 11:59pm.**

**Week 4, 09/25 - 09/29:** Contemporary Misuses of Data: Graphical Manipulation

- Cairo (2019), Introduction, Ch. 1
- Tufte (2001), Ch. 2

**Week 5, 10/02 - 10/06:** Statistical Pitfalls in the Social Science Research

- Head et al. (2015)
- Nelson, Simonsohn, and Simmons (2023)
- Herndon, Ash, and Pollin (2014)

### Module 2: Introduction to Big Data Analysis - Statistical Computing

**Week 6, 10/09 - 10/13:** Introduction to R

- A Gentle Introduction to Tidy Statistics in R (Video)
- Wickham, Çetinkaya-Rundel, and Golemund (2023), Chs. 1 - 9
- **No class Tues. 10/10.**
- **Assignment 2 due Friday, Oct. 13 at 11:59pm.**

**Week 7, 10/16 - 10/20:** Data Visualization

- Wickham, Çetinkaya-Rundel, and Golemund (2023), Chs. 10 - 12

**Week 8, 10/23 - 10/27:** Descriptive Statistics and Working with Data

- Wickham, Çetinkaya-Rundel, and Golemund (2023), Chs. 13 - 15, Ch. 17, Ch. 21

### **Module 3: Big Data and Economics**

**Week 9, 10/30 - 11/03:** Correlation vs. Causation in Economics

- Matthews (2000)
- Card and Krueger (1993)
- Card (1990)
- **Assignment 3 due Friday, Nov. 3 at 11:59pm.**

**Week 10, 11/06 - 11/10:** Case-Study: Race and Inter-generational Mobility

- Einav and Levin (2014)
- Chetty et al. (2018)

**Week 11, 11/13 - 11/17:** Case-Study: Public Policy and Fertility

**Week 12, 11/20 - 11/24:** No class (Thanksgiving break)

- **Assignment 4 due Tuesday, Nov. 21 at 11:59pm**

### **Final Project Presentations**

**Week 13, 11/27 - 12/01:** Wrapping up and Group Presentations

**Week 14, 12/04 - 12/08:** Group Presentations

- **Final Project Report due Friday, Dec. 8 at 11:59pm.**

## References

- Cairo, Alberto. 2019. *How Charts Lie: Getting Smarter about Visual Information*. W. W. Norton & Company, October.
- Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market." Publisher: SAGE Publications Inc, *ILR Review* 43, no. 2 (January): 245–257. Accessed August 12, 2023. <https://doi.org/10.1177/001979399004300205>. <https://doi.org/10.1177/001979399004300205>.
- Card, David, and Alan B. Krueger. 1993. *Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania*. Working Paper, October. Accessed August 4, 2023. <https://doi.org/10.3386/w4509>. <https://www.nber.org/papers/w4509>.
- Chetty, Raj, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2018. *Race and Economic Opportunity in the United States: An Intergenerational Perspective*. Working Paper, March. Accessed August 4, 2023. <https://doi.org/10.3386/w24441>. <https://www.nber.org/papers/w24441>.
- D'Ignazio, Catherine, and Lauren F. Klein. 2020. *Data Feminism*. MIT Press, March.
- Einav, Liran, and Jonathan Levin. 2014. "Economics in the age of big data." Publisher: American Association for the Advancement of Science, *Science* 346, no. 6210 (November): 1243089. Accessed August 4, 2023. <https://doi.org/10.1126/science.1243089>. <https://www.science.org/doi/full/10.1126/science.1243089>.
- Gould, Stephen Jay. 1996. *Mismeasure of Man*. W. W. Norton & Company, February.
- Head, Megan L., Luke Holman, Rob Lanfear, Andrew T. Kahn, and Michael D. Jennions. 2015. "The Extent and Consequences of P-Hacking in Science." Publisher: Public Library of Science, *PLOS Biology* 13, no. 3 (March): e1002106. Accessed August 4, 2023. <https://doi.org/10.1371/journal.pbio.1002106>. <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002106>.
- Herndon, Thomas, Michael Ash, and Robert Pollin. 2014. "Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff." *Cambridge Journal of Economics* 38, no. 2 (March): 257–279. Accessed August 4, 2023. <https://doi.org/10.1093/cje/bet075>. <https://doi.org/10.1093/cje/bet075>.
- Matthews, Robert. 2000. "Storks Deliver Babies ( $p=0.008$ )." *Teaching Statistics* 22 (2): 36–38. Accessed August 4, 2023. <https://doi.org/10.1111/1467-9639.00013>. <https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-9639.00013>.
- Nelson, Leif, Uri Simonsohn, and Joe Simmons. 2023. [109] *Data Falsificada (Part 1): "Clusterfake"*, June. Accessed August 4, 2023. <https://datacolada.org/109>.
- O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.
- Tufte, Edward R. 2001. *The Visual Display of Quantitative Information*. Graphics Press.
- Wickham, Hadley, Mine Çetinkaya-Rundel, and Garrett Grolemund. 2023. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media, Incorporated.

