2022

Cooking Up a Data Literacy Course

Claire Nickerson

Cooking Up a Data Literacy Course

Claire Nickerson, Learning Initiatives and Open Educational Resources Librarian, Forsyth Library, Fort Hays State University, cenickerson@fhsu.edu

NUTRITION INFORMATION

This asynchronous online course, Interdisciplinary Studies 815: Introduction to Data, was developed for graduate students in the information analysis and communication concentration of the Fort Hays State University master of liberal studies degree. The course is designed for professionals who need to make data-driven decisions such as educators, policy makers, and nonprofit employees. It is a survey course, so it does not go into great depth on any of the topics covered but rather provides a basic grounding for developing further data literacy skills. It exclusively uses zero-cost resources, including openly licensed content, library-licensed e-books and articles, and free online content. The course uses weekly mini-assessments as well as three major assignments: a data memo, a data visualization, and a final video presentation.

TARGET AUDIENCE AND NUMBER SERVED

The cap is 15 students. A class size of at least 5 students allows for more peer interaction and feedback. The course was designed for graduate students, but the content would also be appropriate for undergraduates.

LEARNING OBJECTIVES

Students will

- identify different types and sources of data, design a research question that can be answered using data, and determine what type of data is appropriate to answer a given research question
- conduct basic data analysis and visualization and be aware of more advanced tools used in working with data
- articulate the narrative behind data findings and communicate insights about data in various formats to targeted audiences

COOKING TIME

The course has been taught as a 3-credit-hour course over a 15-week semester, but it could be condensed into a trimester or short course.

DIETARY GUIDELINES

The course ties closely into ACRL's *Framework for Information Literacy for Higher Education*. Students learn about Searching as Strategic Exploration and that Authority Is Constructed and Contextual through examining different sources of existing data, such as census and other polling data, and processes for collecting their own data. They participate in Information Creation as a Process and Research as Inquiry by choosing a research question, conducting a survey, and gradually adding background information and audiovisual elements to the presentation of their results throughout the course. Students also read news articles providing case studies about the value of data in the real world.

INGREDIENTS

Reading and video topics by week (see Instructions).

PREPARATION

1. Set up rubrics for grading all assignments. Since the course is asynchronous, scores and comments on rubrics are the primary means by which students receive feedback, so providing detailed comments on each section of the rubric is advised.
2. Review and update news articles and other case studies used as course materials.
3. For each reading, provide reading time estimates using the Niram Read-O-Meter (https://niram.org/read/) and textual complexity estimates using the Lexile Analyzer (https://hub.lexile.com/analyzer).
4. Ask students to complete an intake survey asking the following questions (all optional):
    a. Is English your native language? Would you like feedback on grammar and syntax in your coursework, or only on content?
    b. Are you willing to grant copyright permissions for the instructor to

use your work in specific contexts (e.g., as an example for future classes)?

c. Is there anything you would like the instructor to know about you as a student?

INSTRUCTIONS

Students complete the following readings and assessments/class activities asynchronously. Students receive feedback from the instructor in discussion board postings, via comments on assignment grading rubrics, and during scheduled virtual office hours. Modules are posted at the beginning of the semester, so students can look ahead at assignments and instructions.

Week 1

Topic: What are data? The rise of data capital as a business asset, defining data, and data types.

Readings:

Andrus, Calvin, Jon Cook, and Suresh Sood. "Definitions of Data." Chapter 3 in *Data Science: An Introduction*. Wikibooks, 2017. https://en.wikibooks.org/wiki/Data_Science:_An_Introduction/Definitions_of_Data.

Minitab Blog Editor. "Understanding Qualitative, Quantitative, Attribute, Discrete, and Continuous Data Types." *Minitab Blog*, April 28, 2017. http://blog.minitab.com/blog/understanding-statistics/understanding-qualitative-quantitative-attribute-discrete-and-continuous-data-types.

MIT Technology Review. *The Rise of Data Capital*. White paper, *MIT Technology Review* Custom + Oracle. March 21, 2016. http://files.technologyreview.com/whitepapers/MIT_Oracle+Report-The_Rise_of_Data_Capital.pdf.

Assessment: "What Are Data?" reflection video. Create a video explaining in their own words what data are and why they are valuable. (Students may use any tools they like, but VidGrid is recommended for students who have not created videos before.) Post the video on the discussion board.

Week 2

Topic: Why use data? Data-driven decision-making, news articles about companies making poor decisions based on poor data.

Readings:

Harvard Business Review Analytic Services. *The Evolution of Decision Making: How Leading Organizations Are Adopting a Data-Driven Culture*. Cambridge, MA: Harvard Business School Publishing and SAS, 2012. https://hbr.org/resources/pdfs/tools/17568_HBR_SAS%20Report_webview.pdf.

Pleven, Liam. "Pecan Buyers Shelled by Bad Data." *Wall Street Journal*, April 6, 2012. https://www.wsj.com/articles/SB10001424052702304072004577326002427369224.

Soble, Jonathan. "Fear of More Bad Data Adds to Woes of Kobe Steel." *New York Times*, October 13, 2017. https://www.nytimes.com/2017/10/12/business/kobe-steel-japan-trains.html.

Class activity: Data-driven decision-making discussion.

Assessment: Make a discussion board post about a time when they needed more data to make a decision.

Week 3

Topic: Data ethics. Research bias, guidelines for research with human subjects, HIPPA, and FERPA.

Readings:

Office for Civil Rights. "HIPAA for Professionals." HHS.gov. Last modified May 17, 2021. https://www.hhs.gov/hipaa/for-professionals/index.html.

Rosser, Sue V. "Bias, Research." In *Encyclopedia of Bioethics*, edited by S.G. Post, 1:273–78. New York: Macmillan Reference, November 30, 2003.

Steneck, Nicholas H. "Data Management Practices." In *Introduction to the Responsible Conduct of Research*, 87–102. Washington, DC: Department of Health and Human Services, 2007.

University of Texas at San Antonio Libraries. "What Is FERPA?" YouTube video, 1:22. University of Texas at San Antonio Libraries, May 26, 2011. https://youtu.be/_5XpRGd8O44.

Class activity: Data privacy scenario discussion.

Assessment: Scholarly article bias assessment. Analyze an empirical scholarly article for potential biases. Write a scenario about a researcher, health-care professional, or educator who violates data privacy rules. Then discuss these scenarios with other students using the discussion board.

Week 4

Topic: Searching for data. Overview of census data tools, where to find polling data and public statistics, how to select a research topic, and how to write a research question.

Readings:
New Literacies Alliance. "Ask the Right Questions." January 25, 2017. https://online.newliteraciesalliance.org/courses/course-v1:NLA+IL1013+20_21/about

Nickerson, Claire. "Finding Data," 2018. https://fhsuguides.fhsu.edu/journalism/data.

University of Michigan–Flint. "Select a Topic." 2022. Thompson Library. https://libguides.umflint.edu/research/topic

US Census Bureau. "Data Tools and Apps." Accessed August 16, 2021. https://www.census.gov/data/data-tools.html.

Class activity: Research question.
Assessment: Data set search. Write a research question and locate a data set that provides useful background information related to the topic of the question. (Students usually choose questions related to their personal interests, local communities, or professional lives. Instructor feedback on questions often relates to broadening or narrowing the question and ensuring that it is feasible to collect data that will answer the question.)

Week 5
Topic: Gathering data. Data collection methods, with an emphasis on collecting survey data
Readings:
Blackstone, Amy. "Survey Research: A Quantitative Technique." Chapter 8 in *Sociological Inquiry Principles: Qualitative and Quantitative Methods*. Orono, ME: 2012 Book Archive. https://2012books.lardbucket.org/books/sociological-inquiry-principles-qualitative-and-quantitative-methods/s11-survey-research-a-quantitative.html.

Blackstone, Amy. "Other Methods of Data Collection and Analysis." Chapter 12 in *Sociological Inquiry Principles: Qualitative and Quantitative Methods*. Orono, ME: 2012 Book Archive. https://2012books.lardbucket.org/books/sociological-inquiry-principles-qualitative-and-quantitative-methods/s15-other-methods-of-data-collecti.html.

Kaplowitz, Rella. "How to Collect Data." Data Playbook. Accessed August 16, 2021. https://www.schusterman.org/playbooks/data/how-to-collect-data/.

Assessment: Data collection plan. Write a data collection plan for conducting a survey to answer the research question.

Week 6
Topic: Basic tech skills. Microsoft Excel basics, statistical software options.
Readings:
Microsoft Office. "Excel Video Training." Microsoft Office. Last modified 2022. https://support.office.com/en-us/article/excel-for-windows-training-9bc05390-e94c-46af-a5b3-d7c22f6990bb.

Goldwater, Eva. "Using Excel for Data Analysis—Caveats." University of Massachusetts Amherst. Last modified February 2007. http://people.umass.edu/evagold/excel.html.

Rubin, Denis. "Which Statistical Software to Use?" Quantitative Analysis Guide. NYU Libraries. Last modified May 25, 2022. https://guides.nyu.edu/quant/statsoft.

Assessment: Excel feature video. Pick one tool or feature in Excel and make a video explaining how it works and why it is useful.

Week 7
Topic: Storytelling. How to write about data, examples of news stories centered around data.
Readings:
Gallo, Carmine. "Jeff Bezos Banned PowerPoint in Meetings. His Replacement Is Brilliant." Inc.com. Last modified April 25, 2018. https://www.inc.com/carmine-gallo/jeff-bezos-bans-powerpoint-in-meetings-his-replacement-is-brilliant.html.

Klass, Gary M. "Tabulating the Data and Writing about the Numbers." In *Just Plain Data Analysis: Finding, Presenting, and Interpreting Social Science Data*, 2nd ed., 61–78. Lanham, MD: Rowman & Littlefield, 2012.

Outside in America team. "Bussed Out: How America Moves Thousands of Homeless People around the Country." *Guardian*, December 20, 2017. https://www.theguardian.com/us-news/ng-interactive/2017/dec/20/bussed-out-america-moves-homeless-people-country-study.

Schwirtz, Michael, Michael Winerip, and Robert Gebeloff. "The Scourge of Racial Bias in New York State's Prisons." *New York Times,* December 3, 2016. https://www.nytimes.com/2016/12/03/nyregion/new-york-state-prisons-inmates-racial-bias.html.

United Nations Economic Commission for Europe. *Making Data Meaningful, Part 1: A Guide to Writing Stories about Numbers*. New York

and Geneva: United Nations, 2009. https://unece.org/fileadmin/DAM/stats/documents/writing/MDM_Part1_English.pdf.

Assessment: News article evaluation, Data Memo (Midterm 1). Choose and evaluate a news article for data storytelling skills (details below).

Data Memo

1. Consider how you could answer the research question you wrote last week using survey data.
2. Make a data collection plan for gathering the survey data you need to answer your question. Your data collection plan should be approximately two double-spaced pages and address the following issues:
    a. What resources you will need to implement your plan.
    b. How you will administer your survey and an explanation of why you picked that method (for instance, a survey could be administered verbally, on paper, or using an online tool such as Google Forms).
    c. Who you will include in your survey.
    d. How the data you plan to collect will answer your research question.
3. Implement your data collection plan. Survey data will be due in week 7. If your data collection plan involves the internet, please *include a link to your survey* in your data collection plan submission.

Week 8

Topic: Types of visualizations. Principles of data visualization design and tips for choosing appropriate types of data visualizations.

Readings:

Nelson, Stephen L. "Ten Tips for Visually Analyzing and Presenting Data." Chapter 15 in *Excel Data Analysis for Dummies*, 2nd ed. New York: John Wiley & Son, 2014.

Pierson, Lillian, and Jake Porway. "Following the Principles of Data Visualization Design." Chapter 9 in *Data Science for Dummies.* Somerset, NJ: John Wiley & Sons, 2016.

Assessment: Data visualization scenarios. Given a series of scenarios, describe what kind of data visualization would be most effective for communicating specific data to a specific audience in order to achieve a specific goal and why.

Week 9

Topic: Creating visualizations. When and how to include figures and tables in your writing, tutorials on creating charts in Excel.

Readings:

UNC Chapel Hill Writing Center. "Figures and Charts." Last modified 2018. https://writingcenter.unc.edu/tips-and-tools/figures-and-charts/.

Infobase Cloud Learning. "Excel Advanced: Exploring Charts and Data and Creating Pivot Tables." 2021. https://learningcloud.infobase.com/5972/learnit/144452 (Subscription required)

Microsoft 365. "Insert Excel Data in PowerPoint." Accessed August 16, 2021. https://support.microsoft.com/en-us/office/insert-excel-data-in-powerpoint-19767daf-672c-43bc-bda1-330b242c57c9.

Assessment: Data visualization creation. Choose a data set and create four different types of data visualizations based on that data set, explaining the advantages and disadvantages of each.

Week 10

Topic: Preparing data. Data quality and how to normalize, anonymize, and clean data.

Readings:

Borgatti, Steve. "Normalizing Variables." BA 762: Research Methods. Last modified 2010. http://www.analytictech.com/ba762/handouts/normalization.htm.

Loshin, David. "Data Quality and Information Compliance." In *Business Intelligence: The Savvy Manager's Guide*. San Francisco: Elsevier Science & Technology, 2003.

Nelson, Stephen L. "Scrub-a-Dub-Dub: Cleaning Data." Chapter 3 in *Excel Data Analysis for Dummies*, 2nd ed. New York: John Wiley & Son, 2015.

Roark, Kendall. "Sensitive Research Data Management: Approaches/Techniques for Sharing Sensitive Data." Research guide. Purdue University Libraries. Last modified 2019. https://guides.lib.purdue.edu/c.php?g=352785&p=2377926.

Assessment: Data set evaluation. Evaluate data cleaning and preparation of a user-contributed data set in a public repository.

Week 11

Topic: Types of data analysis. Types of data analysis and what types of questions data can be used to answer

Readings:

Leek, Jeff. "The Data Analytic Question." Chapter 2 in *The Elements of Data Analytic Style: A Guide for People Who Want to Analyze Data*, 3–9. Leanpub, 2015. https://worldpece.org/sites/default/files/datastyle.pdf.

Peng, Roger D., and Elizabeth Matsui. "Types of Questions." Section 3.1 in *The Art of Data Science*. Bookdown, 2017. https://bookdown.org/rdpeng/artofdatascience/types-of-questions.html.

Assessment: Data analysis types, data visualization (Midterm 2; details below).
Describe what type of data analysis they will use to answer their research question and why.

Data Visualization

1. Create one or more data visualizations to accompany your Data Memo from Week 7. Your data visualizations should:
    a. be directed at the same audience to whom you wrote your data memo; and
    b. use the data visualization creation principles you learned in weeks 8 and 9.
2. Write an explanation of why you made the choices you did in creating your data visualizations. Be sure to refer back to the readings.
3. Submit your data visualizations and explanation through Blackboard.

Week 12–13

Topic: Storing and managing data. The research data life cycle, metadata, and metadata standards. How to write a data management plan, version control, and promoting reuse of data.
Readings:
Gilliland, Anne J. "Introduction to Metadata," v. 2.0. Getty Center. Accessed August 16, 2021. https://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html.

Jones, Sarah. "How to Develop a Data Management and Sharing Plan." Data Curation Centre. September 8, 2011. https://www.dcc.ac.uk/guidance/how-guides/develop-data-plan.

Riley, Jenn, and Davin Becker. "Seeing Standards: A Visualization of the Metadata Universe." Scholars Portal Dataverse. Last modified 2018. https://doi.org/10.5683/SP2/UOHPVH.

UK Data Service. "Data Backup." Last modified 2021. https://www.ukdataservice.ac.uk/manage-data/store/backup.aspx.

UK Data Service. "Research Data Lifecycle." Last modified 2019. https://www.ukdataservice.ac.uk/manage-data/lifecycle.

UK Data Service. "Version Control and Authenticity." Last modified 2021. https://www.ukdataservice.ac.uk/manage-data/format/versioning.aspx.

White, Ethan P., Elita Baldridge, Zachary T. Brym, Kenneth J. Locey, Daniel J. McGlinn, and Sarah R. Supp. "Nine Simple Ways to Make It Easier to (Re)Use Your Data." *Ideas in Ecology* 6, no 2 (2013). https://ojs.library.queensu.ca/index.php/IEE/article/view/4608.

Whitmire, Amanda. "Data Management throughout the Research Lifecycle." Figure. Figshare. Last modified August 20, 2013. https://doi.org/10.6084/m9.figshare.774628.v2.

Woodbury, David. "Metadata Tutorial." Vimeo, 2009. https://vimeo.com/3161893.

White, Ethan P., Elita Baldridge, Zachary T. Brym, Kenneth J. Locey, Daniel J. McGlinn, and Sarah R. Supp. "Nine Simple Ways to Make It Easier to (Re)Use Your Data." *Ideas in Ecology and Evolution* 6, no. 2 (August 30, 2013).

Assessment: Research data life cycle diagram, metadata standard explanation, data management plan.
Draw a diagram of the research data life cycle and list at least three steps that might occur during each stage. Choose and explain a metadata standard.
Write a data management plan for their survey data.

Week 14

Topic: Advanced tools. Broad overview of SQL, Python, R, MySQL, Hadoop, Tableau, and others
Readings:
Salehian, Iman. "Tableau Public." In *Introduction to Digital Humanities: Concepts, Methods, and Tutorials for Students and Instructors*, by Johanna Drucker with David Kim, Iman Salehian, and Anthony Bushong, 89–95. Last modified March 12, 2020. https://web.archive.org/web/20200312112540/http://dh101.humanities.ucla.edu/?page_id=163.

Lerner, K. Lee, and Brenda Wilmoth Lerner,

eds. "SQL." In *Computer Sciences*, 2nd ed., 2:238–40. Software and Hardware. Detroit, MI: Macmillan Reference USA, 2013.

Pierson, Lillian, and Jake Porway. "Computing for Data Science." Part 4 in *Data Science for Dummies,* 2nd ed. Somerset, NJ: John Wiley & Sons, 2017.

Pierson, Lillian, and Jake Porway. "Talking About Tableau Public." In *Data Science for Dummies*. Somerset, NJ: John Wiley & Sons, 2017.

Sheikh, Nauman. "Big Data, Hadoop, and Cloud Computing." Chapter 11 in *Implementing Analytics: A Blueprint for Design, Development, and Adoption*. San Francisco: Elsevier Science & Technology, 2013.

"What Is MySQL?" In *MySQL 8.0 Reference Manual*. MySQL, 2021. https://dev.mysql.com/doc/refman/8.0/en/what-is-mysql.html.

Assessment: Advanced data tool overview. Choose and explain the use of an advanced data tool.

Week 15

Topic: Additional topics. Big data and its implications for the future

Readings:

Meyer, Robinson. "The Cambridge Analytica Scandal, in Three Paragraphs." *Atlantic*, March 20, 2018. https://www.theatlantic.com/technology/archive/2018/03/the-cambridge-analytica-scandal-in-three-paragraphs/556046/.

Pence, Harry E. "What Is Big Data and Why Is It Important?" *Journal of Educational Technology Systems* 43, no. 2 (December 1, 2014): 159–71. https://doi.org/10.2190/ET.43.2.d.

Assessment: Data presentation (Final)

Data Presentation—Due Week 15

1. Create a video presentation based on the data set you have been working with all semester. This is the final project for the course, so be sure to incorporate all of the previous concepts, including but not limited to
    a. What the research question is and how your data set answers it
    b The elements of your data collection plan
    c. How you analyzed your data
    d. Telling a story to a specific audience
    e. Choosing appropriate visualization types and following best practices for data visualization design
    f. Your plans for data storage, backup, security, and reuse
2. You may use any tool you like to create the video, but it must include visual aids (with at least one data visualization), text, verbal explanation, and relevant citations. If you are not sure what tools to use, I recommend creating slides in PowerPoint and then recording yourself presenting your slides using the VidGrid recorder (https://app.vidgrid.com).
3. Upload your video here, or provide a link to it.

REVIEWS/ASSESSMENT STRATEGY

Each week of the course includes one or two mini-assessments as described above, such as a discussion board, video creation assignment, reflection paper, or quiz. However, the three major assessments for the course all focus on methods of using data to communicate effectively. Students administer a survey on a topic of their choice and then share the results as a memo in week 5, a data visualization in week 10, and a video presentation in week 15, adding background information related to their topic along the way.

ADAPTING THE RECIPE

After having offered the course twice, I would suggest making some modifications.

- The major assignment to conduct a survey could be expanded to give students additional data collection options, such as focus groups or interviews.
- The final video presentations students submitted would have benefitted from a time limit to encourage concision.
- One of the shorter assessments was to create instructional videos explaining specific features of Excel, which were submitted directly to the instructor but would be more beneficial if collected in a library where other students could see them.
- Regarding course materials, students had difficulty identifying biases in empirical research articles and might benefit from more in-depth coverage and examples.
- They also submitted data visualizations that used color to convey meaning without providing additional data labels, suggesting that at least one reading about how to create accessible visualizations should be added.

Student feedback on the course has generally been positive. In course evaluations, one student said that the course "was well organized, balanced, clearly communicated and provided ample opportunity to tailor learning to real-life situations," and another student said that it "allowed for flexibility on deadlines while encouraging a manageable pace." The primary criticism of the course was that the content level was too low for students who had taken other data-focused courses first or were further along in their careers.

ADDITIONAL RESOURCES

Syllabus available at https://drive.google.com/file/d/17Ad-BexYUTUEoI7pmaB10xUf6L-GLaMPB/view?usp=sharing.