

---

## Term Information

Effective Term Autumn 2022

## General Information

Course Bulletin Listing/Subject Area English  
Fiscal Unit/Academic Org English - D0537  
College/Academic Group Arts and Sciences  
Level/Career Undergraduate  
Course Number/Catalog 2200  
Course Title Data Analysis for the Digital Humanities  
Transcript Abbreviation DataAnalysisDigHum  
Course Description In this course students will learn principles of data analysis applicable to English Studies in the digital humanities. Methods of the digital humanities, which apply tools of statistical, computational, and spatial analysis to humanistic topics like rhetoric and literary studies, offers the ideal vehicle for introducing students in the liberal arts to quantitative techniques for data analysis.  
Semester Credit Hours/Units Fixed: 3

## Offering Information

Length Of Course 14 Week, 12 Week, 8 Week, 7 Week, 6 Week, 4 Week  
Flexibly Scheduled Course Never  
Does any section of this course have a distance education component? No  
Grading Basis Letter Grade  
Repeatable No  
Course Components Lecture  
Grade Roster Component Lecture  
Credit Available by Exam No  
Admission Condition Course No  
Off Campus Never  
Campus of Offering Columbus, Lima, Mansfield, Marion, Newark

## Prerequisites and Exclusions

Prerequisites/Corequisites Prerequisite: English 1110  
Exclusions  
Electronically Enforced Yes

## Cross-Listings

Cross-Listings

## Subject/CIP Code

Subject/CIP Code 09.0702  
Subsidy Level Baccalaureate Course  
Intended Rank Freshman, Sophomore, Junior

## Requirement/Elective Designation

Mathematical and Quantitative Reasoning (or Data Analysis)

## Course Details

### Course goals or learning objectives/outcomes

- Students learn the history of the digital humanities, understand methods of computational analysis of digital humanities data such as literary and persuasive texts, understand and apply principles of ethical qualitative research with human subjects.

### Content Topic List

- What is data analysis and the digital humanities?
  - Introduction to R and statistical thinking
  - Statistics and qualitative data
  - What is humanities data?
  - Assembling a dataset
  - Corpus data
  - Analyzing humanities data
  - Visualizing data sets
  - Working with dates in R
  - Making arguments through time
  - Geocoding 1: Spatial analysis
  - Geocoding 2: Making arguments with maps
- No

### Sought Concurrence

## Attachments

- Data Analysis.pdf: Proposal and Syllabus  
*(Syllabus. Owner: Lowry, Debra Susan)*
- GE Submission Form, Data Analysis for DH, English 2200.pdf: GE Submission Form  
*(Other Supporting Documentation. Owner: Lowry, Debra Susan)*

## Comments

## Workflow Information

Status	User(s)	Date/Time	Step
Submitted	Lowry, Debra Susan	03/15/2021 05:16 PM	Submitted for Approval
Approved	Winstead, Karen Anne	03/16/2021 08:05 AM	Unit Approval
Approved	Vankeerbergen, Bernadette Chantal	03/16/2021 05:14 PM	College Approval
Pending Approval	Jenkins, Mary Ellen Bigler Hanlin, Deborah Kay Oldroyd, Shelby Quinn Hilty, Michael Vankeerbergen, Bernadette Chantal	03/16/2021 05:14 PM	ASCCAO Approval

# English 2200: Data Analysis for the Digital Humanities

## General Education Course Proposal

### *Proposal Information*

---

- **Proposed Number:** 2200xx
- **Proposed Title:** Data Analysis for the Digital Humanities
- **Proposed GE Category:** Mathematical and Quantitative or Data Analysis (Foundations)
- **Proposed Embedded Literacies Category:** Data Analytics
- **Proposal Type:** New
- **Proposal prepared by:** John Jones

### *Description, Rationale, and Links to Other English Classes*

---

In this course, students will learn principles of data analysis applicable to English Studies in the digital humanities. The methods of the digital humanities, which apply tools of statistical, computational, and spatial analysis to humanistic topics like rhetoric and literary studies, offers the ideal vehicle for introducing students in the liberal arts to quantitative techniques for the analysis of data. As a methods course, students will be introduced to a range of data analysis practices in the digital humanities and apply these methods to research projects of their choosing.

In this class, students will...

- Learn the history of the digital humanities—from early explorations of humanistic topics using digital tools to the modern emergence of the field—and a range of computational, spatial, and statistical methods associated with the analysis of humanistic topics and data.
- Identify and analyze research topics and questions
- Understand methods of computational analysis of DH data such as literary and persuasive texts.
- Understand methods of corpus analysis to large textual data sets.
- Understand principles of ethical qualitative research with human subjects, including interview and survey data and data collected from public sources like social data.
- Understand basic statistical analysis of qualitative and quantitative data sets.
- Understand methods of spatial data collection and argument.
- Apply methods of computational, statistical, and spatial data analysis to research questions in the digital humanities and craft visual and written arguments in response to these questions.

In addition to meeting the Data Analysis and Data Analytics GE categories, this course will provide students in literary and culture studies as well as WRL an opportunity to familiarize themselves with a range of methods related to digital research, preparing them to apply these methods throughout the literature and WRL curriculum.

The course complements the offerings for the Integrated Major in Mathematics and English (IMME). Like that program, Data Analytics for the Digital Humanities introduces students to quantitative and

qualitative methodologies that are appropriate for the study of humanistic disciplines like English Studies. In this way, the course provides a snapshot of the IMME program for non-IMME majors and demonstrates that our curriculum is invested in addressing a range of methods for the study of rhetorical, literary, and cultural texts.

**Relevance to the GE Category**

This class will effectively address the expected learning outcome for the Mathematical and Quantitative or Data Analysis (Foundations) category.

Foundations: Mathematical and Quantitative or Data Analysis		
Goals	Expected Learning Outcomes	Related Course Content
<p><b>Goal: Successful students will be able to apply quantitative or logical reasoning and/or mathematical/statistical analysis methodologies to understand and solve problems and to communicate results.</b></p>	<p><b>Successful students are able to ...</b></p> <p><b>1.1</b> Use logical, mathematical and/or statistical concepts and methods to represent real-world situations.</p>	<p><b>In this course, students will ...</b></p> <p>1.1 Use logical and statistical concepts to represent humanistic research topics</p>
	<p><b>1.2</b> Use diverse logical, mathematical and/or statistical approaches, technologies, and tools to communicate about data symbolically, visually, numerically, and verbally.</p>	<p>1.2 Use mathematical and statistical approaches and software to create data visualizations of digital humanities subjects</p>
	<p><b>1.3</b> Draw appropriate inferences from data based on quantitative analysis and/or logical reasoning.</p>	<p>1.3 Understand how to draw appropriate inferences from the analysis of DH data sources</p>
	<p><b>1.4</b> Make and evaluate important assumptions in estimation, modeling, logical argumentation, and/or data analysis.</p>	<p>1.4 Make and evaluate assumptions in logical argumentation and analysis of data sets produced through DH methods</p>
	<p><b>1.5</b> Evaluate social and ethical implications in mathematical and quantitative reasoning.</p>	<p>1.5 Understand and evaluate ethical uses of quantitative reasoning DH research.</p>

This class will also meet the ELOs for the Data Analytics category of embedded literacy.

Data Analysis Literacy		
Goal	Expected Learning Outcomes	Related Course Content
<p>Successful students will meet the goals for <i>either</i> a Quantitative Data Analysis (A) or Qualitative Data Analysis (B) course.</p> <p><b>Quantitative Data Analysis (A) Goal:</b> Successful students develop skills in drawing conclusions and critically evaluating results based on data.</p>	<p>Successful students are able to ...</p> <p><b>1.1A</b> explain basic concepts of statistics and probability.</p>	<p>In this course, students will ...</p>
	<p><b>1.2A</b> apply methods needed to analyze and critically evaluate statistical arguments.</p>	
	<p><b>1.3A</b> recognize the importance of statistical ideas.</p>	
	<p><b>1.4A</b> evaluate the social and ethical implications of data collection and analysis, especially in relation to human subjects.</p>	
<p><b>Qualitative Data Analysis (B) Goal:</b> Successful students develop skills in drawing conclusions and critically evaluating results based on data.</p>	<p><b>1.1B</b> explain the utility of different approaches to qualitative data analysis.</p>	<p>1.1B Explain the utility of spatial and statistical methods to qualitative analysis of DH data</p>
	<p><b>1.2B</b> apply key methods and tools in qualitative data analysis.</p>	<p>1.2B Apply spatial and statistical methods for the analysis of DH data sets</p>
	<p><b>1.3B</b> interpret the results of qualitative data analysis to answer research question(s).</p>	<p>1.3B Be able to interpret data findings from their qualitative research projects to answer specific research questions</p>
	<p><b>1.4B</b> evaluate the social and ethical implications of data collection and analysis, especially in relation to human subjects.</p>	<p>1.4B Be able to conduct ethical data collection and analysis of human subject data</p>

**Staffing**

Susan Lang, Scott DeWitt, and John Jones are all interested in teaching this class.

***Syllabus Status***

---

Attached.

# Data Analysis for the Digital Humanities

English 2200

**Instructor:** John Jones, Associate Professor

**Email:** jones.6181@osu.edu

**Office hours:** Tuesdays, 11a–12p, and by appointment

## About this course

---

### Course description

In this course, students will study rhetorical, literary, and cultural texts and related data sets using the tools and methods of the digital humanities (DH). Where English Studies traditionally uses methods such as *rhetorical analysis* or *close reading* to examine individual or small groups of texts, English scholars in the digital humanities use techniques such as statistical and computational analysis to examine large data sets consisting of hundreds or thousands of texts and visualize this data in ways that highlight not only logical and textual relationships but spatial ones as well. These methods of *distant reading*<sup>1</sup> are made possible by the rapid scanning, analysis, and visualization of data enabled by digital computation, and they are used to complement and extend the traditional methodologies of English Studies in the understanding of reading, writing, and culture.

As a methods course, students will be introduced to a range of data analysis practices in the digital humanities and apply these methods to research projects of their choosing. In doing so, they will learn principles of statistical reasoning and analysis, introductory programming concepts for the computational analysis of texts, and methods for visualizing these results.

Students are not required to have a background in statistics or programming to take the course. This course teaches students all the skills in introductory statistics and programming necessary for the application of digital humanities methodologies to English Studies.

In this class, students will

- Learn the history of the digital humanities—from early explorations of humanistic topics using digital tools to the modern emergence of the field—and a range of computational, spatial, and statistical methods associated with the analysis of humanistic topics and data.
- Understand methods of computational analysis of digital humanities data such as literary and persuasive texts.
- Identify and analyze research topics and questions in the humanities that can be answered with statistical or computational methods
- Understand methods of corpus analysis to large textual data sets.
- Understand principles of ethical qualitative research with human subjects, including interview and survey data and data collected from public sources like social data.
- Understand basic statistical analysis of qualitative and quantitative data sets.
- Understand methods of spatial data collection and argument.
- Apply methods of computational, statistical, and spatial data analysis to research questions in the digital humanities and craft visual and written arguments in response to these questions.

In addition to satisfying the Mathematical and Quantitative or Data Analysis (Foundations) category and the data analysis embedded literacy requirements, this

---

<sup>1</sup> Ramsay, S. (2011). *Reading machines: Toward an algorithmic criticism*. Urbana: University of Illinois Press.

course will provide students in the Literary and Culture Studies and Writing, Rhetoric, and Literacy (WRL) concentrations an opportunity to familiarize themselves with a range of methods related to digital humanities research, preparing them to apply these methods throughout the literature and WRL curriculum. While the course will focus on research topics and data sets appropriate to English Studies, the course methods will be applicable to digital humanities work in related humanities fields such as history, philosophy, and religion studies.

## Required texts & materials

We will read articles and selections from books such as:

- Arnold, T., & Tilton, L. (2015). *Humanities data in R: Exploring networks, geospatial data, images, and text*. Springer.
- Canning, J. (2014). *Statistics for the Humanities*. Brighton, UK.
- Gold, M. K., & Klein, L. F. (Eds.). (2019). *Debates in the digital humanities 2019*. University of Minnesota Press. <https://dhdebates.gc.cuny.edu/>
- Jockers, M. L., & Thalken, R. (2020). *Text analysis with R: For students of literature*. Springer.
- Mullen, L. (2015). *Computational historical thinking: With applications in R*. <https://dh-r.lincolnmullen.com/>
- Wickham, H., & Grolemund, G. (2016). *R for data science: Import, tidy, transform, visualize, and model data*. Sebastopol, CA: O'Reilly. <https://r4ds.had.co.nz/index.html>

You will also need to have access to the following

- An OSU email account
- Access to a computer to run R. R can be downloaded at <https://www.r-project.org/>

## Goals and learning outcomes

---

This class will effectively address the expected learning outcome for the Mathematical and Quantitative or Data Analysis (Foundations) category. Successful students

If you do not have regular access to a computer, please contact me as soon as possible and we can discuss alternative options.

You should prepare a back-up plan in case technology mishaps occur. Regularly backup course files (you can do this via Microsoft OneDrive) and make sure that you will have access to your files and know how to access an alternate computer in case yours becomes unavailable.

## Academic misconduct

It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term “academic misconduct” includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct <http://studentlife.osu.edu/csc/>.

## Disability services

The University strives to make all learning experiences as accessible as possible. If you anticipate or experience academic barriers based on your disability (including mental health, chronic or temporary medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, I may request that you register with Student Life Disability Services. After registration, make arrangements with me as soon as possible to discuss your accommodations so that they may be implemented in a timely fashion. SLDS contact information: [slds@osu.edu](mailto:slds@osu.edu); 614-292-3307; [slds.osu.edu](http://slds.osu.edu); 098 Baker Hall, 113 W. 12th Avenue.

will be able to apply quantitative or logical reasoning and/or mathematical/statistical analysis methodologies



to understand and solve problems and to communicate results.

In this course, students will ...

- Use statistical concepts to analyze research topics in English Studies and the digital humanities
- Use mathematical and statistical approaches and software to create visualizations of data
- Understand how to draw appropriate inferences from the analysis of digital humanities data sources
- Make and evaluate assumptions in logical argumentation and analysis of data sets produced through digital humanities methods
- Understand and evaluate ethical uses of quantitative reasoning in digital humanities research.

This class will also meet the ELOs for the Data Analytics category of embedded literacy. Successful students will develop skills in drawing conclusions and critically evaluating results based on data.

- Explain the utility of spatial and statistical methods to qualitative analysis of digital humanities data
- Apply spatial and statistical methods for the analysis of digital humanities data sets
- Be able to interpret data findings from their qualitative research projects to answer specific research questions
- Be able to conduct ethical data collection and analysis of human subject data

## Grading and assignments

---

### Grading

The following grade ranges will be used in the course. Qualitative descriptions of letter grades can be found in section 3335-8-21 of the OSU Trustees' University Faculty Rules.<sup>2</sup>

A:	93–100	C+:	77–79	E:	0–59
A-:	90–92	C:	73–76		
B+:	87–89	C-:	70–72		
B:	83–86	D+:	67–69		
B-:	80–82	D:	60–66		

You will be graded on the following assignments

Quizzes	25%
Workshops & Discussion	25%
Group research project	30%
Final Exam	20%

### List of assignments and grade percentages

### Quizzes

You will take four quizzes throughout the semester covering the mathematical and coding concepts in the course. All of the quizzes will be open book and open notes, but you may not collaborate with your classmates on the quizzes.

### Workshops & discussion

Students will engage in weekly or bi-weekly workshops in which they explore DH methodologies like statistical analysis following exercises in our textbooks or that are provided by me.

Additionally, you will participate in 5 discussion board conversations addressing topics like the development of the digital humanities and the applicability of humanistic research to contemporary issues like big data analysis, machine learning, and algorithmic decision-making.

### Group research projects

Over the course of the semester, students will work in groups to develop plan and complete a project that applies one or more of the DH methodologies introduced in the workshops to explore research questions in the humanities. This project will be completed in stages, as groups will identify a research question or questions, compile a data set, analyze that data using methodologies introduced in the course, and write up their findings.

---

<sup>2</sup> <https://trustees.osu.edu/rules/university-rules/chapter-3335-8-instruction.html>

# Final exam

You will complete a final exam that covers the mathematical and methodological questions in the course. The

exam will include questions from the quizzes as well as new questions that have not appeared on the quizzes. The exam will be open book and open notes, but you may not collaborate with your classmates.

# Schedule

This schedule provides an overview of the course and the sequence of activities and assignments. **Full details about activities, readings, and assignments, along**

**with specific due dates/times are available on Carmen.** Items on the schedule are subject to change with advance notification.

	POTENTIAL TOPICS/ READINGS	ASSIGNMENTS
WEEK 1	What is data analysis and the digital humanities? <ul style="list-style-type: none"> <li>• Gold, "Introduction," <i>Debates in the Digital Humanities</i></li> <li>• Stephen Ramsay, "Humane Computation," <i>Debates in the Digital Humanities 2016</i></li> <li>• Cottom, "More Scale, More Questions: Observations from Sociology," <i>Debates in the Digital Humanities 2016</i></li> </ul>	
WEEK 2	Introduction to R and statistical thinking <ul style="list-style-type: none"> <li>• Mullen, "Introduction," "Installing R and Packages," and "An R Primer," <i>Computational Historical Thinking</i></li> <li>• Arnold and Tilton, "New Data? The Role of Statistics in DH," <i>Debates in the Digital Humanities 2019</i></li> </ul>	Discussion Board 1
WEEK 3	Statistics and qualitative data <ul style="list-style-type: none"> <li>• Canning, Chs 1–3, <i>Statistics for the Humanities</i></li> <li>• Gladstone, "Teaching Quantitative Methods: What Makes It Hard (in Literary Studies)" <i>Debates in the Digital Humanities 2019</i></li> <li>• Geisler and Swarts, "An Introduction to Coding Streams of Language," <i>Coding Streams of Language</i></li> </ul>	Quiz 1
WEEK 4	What is humanities data? <ul style="list-style-type: none"> <li>• Geisler and Swarts, "Designing the Analysis," <i>Coding Streams of Language</i></li> <li>• Grolemond and Wickham, "Section 1: Explore," <i>R for Data Science</i></li> </ul>	
WEEK 5	Assembling a dataset <ul style="list-style-type: none"> <li>• Canning, Chs 4–7, <i>Statistics for the Humanities</i></li> <li>• Jessica Marie Johnson, "Markup Bodies: Black [Life] Studies and Slavery [Death] Studies at the Digital Crossroads," <i>Social Text</i></li> <li>• Geisler and Swarts, "Segmenting the Data" and "Coding the Data," <i>Coding Streams of Language</i></li> </ul>	Quiz 2
WEEK 6	Corpus data <ul style="list-style-type: none"> <li>• Kim, "Building Pleasure and the Digital Archive," <i>Bodies of Information</i></li> <li>• Grolemond and Wickham, "Section 2: Wrangle," <i>R for Data Science</i></li> </ul>	Discussion Board 2

	POTENTIAL TOPICS/ READINGS	ASSIGNMENTS
WEEK 7	Analyzing humanities data <ul style="list-style-type: none"> <li>• Canning, Chs 10–11, <i>Statistics for the Humanities</i></li> <li>• Grolemund and Wickham, "Section 3: Program," <i>R for Data Science</i></li> </ul>	
WEEK 8	Analyzing humanities data <ul style="list-style-type: none"> <li>• Noble, "Toward a Critical Black Digital Humanities," <i>Debates in the Digital Humanities 2019</i></li> <li>• We will work on Textual data set in class.</li> </ul>	Quiz 3
WEEK 9	Visualizing data sets <ul style="list-style-type: none"> <li>• Grolemund and Wickham, "Section 5: Communicate," <i>R for Data Science</i></li> <li>• Canning, Ch 20, <i>Statistics for the Humanities</i></li> </ul>	Discussion Board 3
WEEK 10	Working with dates in R <ul style="list-style-type: none"> <li>• Mullen, "Dates," <i>Computational Historical Thinking</i></li> </ul>	
WEEK 11	Making arguments through time No scheduled readings. Work on Dates sample data set	Group Research Project Proposals due
WEEK 12	Geocoding 1: Spatial analysis <ul style="list-style-type: none"> <li>• Mullen, "Geocoding," <i>Computational Historical Thinking</i></li> <li>• Canning, Ch 16, <i>Statistics for the Humanities</i></li> </ul>	Discussion Board 4
WEEK 13	Geocoding 2: Making arguments with maps <ul style="list-style-type: none"> <li>• <a href="#">Mapping the Republic of Letters</a></li> <li>• Be prepared to work on geocoding sample data set</li> </ul>	Quiz 4
WEEK 14	Workshop: Data analysis for Research Projects No scheduled readings. Bring preliminary analysis of Group Research Projects to class for workshopping and peer review.	
WEEK 15	Library week – Meet with groups to work on Research Project	Discussion Board 5
FINALS WEEK		Final Exam and Group Research Project due