**DIGHUM101 – Practicing the Digital Humanities**

Dr. Evan Muzzall [evan.muzzall@berkeley.edu](mailto:evan.muzzall@berkeley.edu)

**Dates:** Mondays, Tuesdays, Wednesdays, and Thursdays (May 26 – July 2, 2020)

**Time:** 1:00 – 3:00 PM

**Where:** Zoom and GitHub links forthcoming

**Office Hours:** Friday drop-in (hours TBD); by appointment

**Prerequisites:** None!

**Requirements:** Computer, Zoom account, and Internet connection.

**Learning objectives:**

The goal of this course is to teach you basic principles for conducting professional research in the Digital Humanities. You will learn how to program Python in Jupyter Notebooks to access, explore, visualize, and analyze data in humanities contexts. You are strongly encouraged to concurrently enroll in [DIGHUM 100 – Theory and Methods in the Digital Humanities](https://summerdigitalhumanities.berkeley.edu/courses) to compliment this experience with a strong theoretical foundation. By the end of this course, you will learn:

* A variety of Pythonic approaches to explore questions in the humanities
* How to understand data more holistically; its generative process and life cycle
* Strategies for organizing research projects based on your interests
* Methods for data acquisition and visualization, computational text analysis, and machine learning
* The importance Why it is important to develop a more critical lens of your field of study

**Software installations (please have this complete before first day)**

* Download and install Python Anaconda 3.7 distribution: <https://www.anaconda.com/distribution/>
* Windows users only must install Git (Mac users do not have to): <https://git-scm.com/downloads>

**Course materials:**

* Course materials and instructions will be hosted on the course GitHub repository:   
  (instructions forthcoming)

**Assignments and Grading**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Task** | **Assignment** | **Topic** | **Due** | **% of grade** | **Type** |
| Attendance | 1 | (required) | Everyday | 10 | Participation |
| Homework | 2 | Video | July 2 | 20 | Digital format |
| Group project | 3 | Topic | June 18 | 10 | One-page summary |
|  | 4 | Presentation | June 29-30 | 20 | Slideshow |
| Individual project | 5 | Topic | June 18 | 10 | Two-page summary |
|  | 6 | Presentation | July 1-2 | 20 | Jupyter Notebook |
|  | 7 | Repository | July 2 | 10 | GitHub |

NOTE: Grading rubrics will be added to bCourses.

**Assignment descriptions**

1. **Attendance (Sign-in via Zoom).** This course is synchronous and attendance is required for all students. Asynchronous exceptions will be considered on a case-by-case basis, such as for those living in distant time zones. Each class is divided into two types:
   1. **Lecture days:** The instructor will lecture for 60 minutes to explain the day’s topic. The remainder of the time will be for you to work through specific challenge set problems, ask questions, and think about how the topic relates to your projects.
   2. **Project days:** These days will be exclusively focused on your individual and group projects.
2. **Homework: Video (submit to bCourses)**

Record a 3-5 minute video that demonstrates thoughtful reflection about how your ideas regarding digital humanities research grew throughout this course. Use specific examples and perspectives from your individual and group projects. You must appear in at least 30 seconds of footage.

**Group Project.** The group project will require you to work in small groups, read some theoretical articles, and give a presentation to your classmates. It consists of two assignments: a one-page summary (single-spaced) about your group’s theoretical topic and a presentation.

1. **Group Project – Topic (submit one-page single-spaced proposal to bCourses – only one per group)**

Each student will be assigned to a group. Each group must choose a theoretical topic to present on using the articles in the “selected readings” folder on bCourses. Everyone must read three articles from this folder. Each student in a group can choose the same three articles, or each student can choose a different three articles but the topic must be unified and coherent across all group members and readings. Do not do any programming/coding for this group project.

1. **Group Project – Presentation (publish in your GitHub repository):**

Each group will present a 10-12 minute presentation of their topic using PowerPoint, Prezi, .html, pdf, etc. Be creative here! Emphasize visuals, infographics, video, audio, polls, interactivity, etc. so long as it does not detract from your point (and works via Zoom screen share). Keep your talking points clear and concise and minimize the amount of text you use.

**Individual Project.** The purpose of the individual project is to demonstrate your programming competencies in a digital humanities research context. It consists of three parts:

1. **Individual Project – Topic (submit two-page single-spaced prospectus to bCourses):**

The topic for this individual project is your choice and the only limiting factors are your abilities along with the scope and scale of your idea and the data. Think big, but then carve out a little slice of that big idea to focus on for this six-week course.

1. **Individual Project – Presentation (publish in your GitHub repository):**

You are required to present a 5-minute presentation of your individual project. A functioning Jupyter Notebook is the only acceptable format.

1. **Individual Project – GitHub repository (submit URL to bCourses):**

Publish your group and individual presentations to your GitHub repository. Include a README.md file that includes a: 1) title, 2) header image, 3) abstract of your group project, and 4) abstract of your individual project. Group members cannot use the same abstract – these abstracts must be written individually and include reflection about the collaborative process.

**Schedule**

|  |  |  |
| --- | --- | --- |
| **Week** | **Date** | **Topic** |
| 1 | May 26 | Syllabus, example projects, ethics, and Jupyter Notebooks |
|  | May 27 | Python basics: Build your programming vocabulary |
|  | May 28 | Project day: Group assignments; individual topic brainstorm |
|  |  |  |
| 2 | June 1 | Pandas data frames; Data formats (txt, csv, xml, json) |
|  | June 2 | Text data sources (Project Gutenberg and HTRC); Text preprocessing |
|  | June 3 | Data visualization: Histogram, barchart, boxplot, scatterplot, network |
|  | June 4 | Project day: Individual project dataset exploration; Group project questions |
|  |  |  |
| 3 | June 8 | Introduction to geospatial data; GeoPandas |
|  | June 9 | APIs: praw, tweepy |
|  | June 10 | Webscraping |
|  | June 11 | Example project: Walkthrough |
|  |  |  |
| 4 | June 15 | Introduction to Bash and GitHub |
|  | June 16 | Machine learning: Jargons, preprocessing, regression, classification |
|  | June 17 | Machine learning: Document encoding; CountVectorizer and TfidfVectorizer |
|  | June 18 | Project day: Group project topics due; Individual topic and dataset due |
|  |  |  |
| 5 | June 22 | Machine learning: Topic modeling; Word2Vec, BERT |
|  | June 23 | Machine learning: Sentiment analysis |
|  | June 24 | Introduction to image processing and artificial neural networks |
|  | June 25 | Project day: Individual project focus |
|  |  |  |
| 6 | June 29 | Group presentations |
|  | June 30 | Group presentations |
|  | July 1 | Individual presentations |
|  | July 2 | Individual presentations |

**Legal stuff – read this carefully!**

Remember that you are bound to various codes of conduct. See the legal document in “Files” on bCourses. By enrolling in this class, you take full responsibility for your learning.