**DIGHUM101 – Practicing the Digital Humanities**

Dr. Tom van Nuenen

tomvannuenen@gmail.com

**Dates:** Mondays, Tuesdays, Wednesdays, and Thursdays (May 23 – July 1, 2022)

**Time:** 1:00 – 3:00 PM

**Where:** Zoom

**Office Hours:** By appointment on Zoom

**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 to complement 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 lifecycle;
* Strategies for organizing research projects based on your interests;
* Methods for data acquisition and visualization, computational text analysis, and machine learning;
* The importance of developing a critical lens in your field of study.

**SOFTWARE INSTALLATIONS**

***\*Please have this complete before first day\****

* Download and install Python Anaconda 3.9 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 are on bCourses and the course Github repository: https://github.com/dlab-berkeley/DIGHUM101-2022

**ASSIGNMENTS AND GRADING**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Task** | **Assignment** | **Topic** | **Due** | **% grade** | **Points** | **Type** |
| Attendance | 1 | (required) | Everyday | 10 | 20 | Participation |
| Homework | 2 | Video | July 1 | 20 | 40 | Digital format |
| Group project | 3 | Topic | June 20 | 10 | 20 | One-page summary |
|  | 4 | Presentation | June 27-28 | 20 | 40 | Slideshow |
| Individual project | 5 | Topic |  | 10 | 20 | Two-page summary |
|  | 6 | Presentation |  | 20 | 40 | Jupyter Notebook |
|  | 7 | Repository |  | 10 | 20 | GitHub |

Grading is straight scale: A=90-100; B=80-89; C=70-79; D=60-69; F<60

**Assignment descriptions**

1. **Attendance (sign-in via Zoom).** This course is synchronous and attendance is required. 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 take you through the Jupyter notebooks. 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 on 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. Deadline: 1 July 2022

***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 a one-page single-spaced proposal on bCourses – one per group).** Each student will be assigned to a group. Each group must choose a theoretical or critical 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 readings, or each student can choose different readings, but the topic must be unified and coherent across all group members and readings. Where relevant, use the readings to reflect on your individual project work, such as the kinds of approaches and methods you are engaging in. *Do not do any programming/coding for this group project*. Use 10-12 point font and include references cited. Deadline: 20 June 2022
2. **Group Project – Presentation.** 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. Presentations are held in week 6.

***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 a two-page single-spaced prospectus on 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. Use 10-12 point font. Deadline: 20 June 2022
2. **Individual Project – Presentation.** You are required to give a 5-minute presentation of your individual project. A functioning Jupyter Notebook is the only acceptable format. Walk us through the notebook, your findings, and the choices you made in the research process. Presentations are held in week 6.
3. **Individual Project – GitHub repository (submit URL on bCourses).** Publish your individual project as a Jupyter notebook to your GitHub repository. Include a README.md file that includes a: 1) title, 2) header image, and 3) abstract of your research.

Deadline: 1 July 2022

**Note:** you can find example projects for both individual and group projects in the “Example projects” folder on bCourses. Please brown through them to get your mind going on what the possibilities are and what you would like to do.

**SCHEDULE**

|  |  |  |
| --- | --- | --- |
| **Week** | **Date** | **Topics** |
| 1 | May 23 | Syllabus, example projects, and Jupyter Notebooks |
|  | May 24 | Python basics: Build your programming vocabulary |
|  | May 25 | Pandas data frames; Data formats (txt, csv, xml, json) |
|  | May 26 | Project day: Group assignments; individual topic brainstorm |
|  |  |  |
| 2 | May 30 | *No Class: Memorial Day* |
|  | May 31 | Text data sources (Project Gutenberg and HTRC); Text preprocessing |
|  | June 1 | Data visualization: Histogram, barchart, boxplot, scatterplot, network |
|  | June 2 | Project day: Group project questions; Individual project dataset exploration |
|  |  |  |
| 3 | June 6 | Webscraping (BeautifulSoup) |
|  | June 7 | APIs: Google Books, Reddit (praw), Twitter (tweepy) |
|  | June 8 | Network analysis; geospatial data with (GeoPandas) |
|  | June 9 | Example project: Walkthrough |
|  |  |  |
| 4 | June 13 | Introduction to Bash and GitHub |
|  | June 14 | Machine learning: Jargons, preprocessing, regression, classification |
|  | June 15 | Machine learning: Document encoding (CountVectorizer, TfidfVectorizer) |
|  | June 16 | Project day: Group project topics due; Individual topic and dataset due |
|  |  |  |
| 5 | June 20 | Machine learning: Topic modeling |
|  | June 21 | Machine learning: Word2Vec |
|  | June 22 | Machine learning: Classification |
|  | June 23 | Project day: Individual project focus |
|  |  |  |
| 6 | June 27 | Group presentations |
|  | June 28 | Group presentations |
|  | June 29 | Individual Presentations |
|  | June 30 | 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.