2019 HAWAII UNIVERSITY INTERNATIONAL CONFERENCES
ARTS, HUMANITIES, SOCIAL SCIENCES & EDUCATION JANUARY 3 - 5, 2019
PRINCE WAIKIKI HOTEL, HONOLULU, HAWAII

TEACHING SOCIAL, POLITICAL AND ECONOMIC TOPICS TO STUDENT DATA SCIENTISTS: THE CASE OF A DATA SCIENCE COURSE

CHONG, RYLAN
DIVISION OF NATURAL SCIENCE AND MATHEMATICS
CHAMINADE UNIVERSITY
HONOLULU, HAWAII

Dr. Rylan Chong
Division of Natural Science and Mathematics
Chaminade University
Honolulu, Hawaii

Teaching Social, Political, and Economic Topics to Student Data Scientists: The Case of a Data Science Course

Synopsis:

There is a calling to teach social, political, and economic topics in STEM-oriented curriculum. This paper used a case study approach to discuss the integration of the topics in a introductory data science course that is one of six courses of a certificate that will be offered by Chaminade University. The integration approach was to use the topics as subtexts with two subtopic areas of surveying data science, and during the interpretation and evaluation of the results.

Teaching Social, Political, and Economic Topics to Student Data Scientists: The Case of a Data Science Course

Introduction

Students progressing through and graduating from a science, technology, engineering, and math (STEM) program are expected to be aware of the growing significance and impact of one's work and role has upon society. In other words, along with mastering the STEM technical knowledge and skills, it is being able to think beyond ones discipline, and to consider other factors and dimensions (Grasso & Martinelli, 2008) of social, political, and economics. For instance, National Research Council of the National Academies (NRCNA) suggested that science students need to appreciate "the complexity of the policy world, as well as an understanding of the assumptions underlying divergent policy framings, experts judgements, consensus building techniques, and standard analytic methods and approaches" (National Research Council of the National Academies, 2012, p. 6). The NRCNA continues, students need to "recognize the limits of the persuasive power of scientific reasoning, the substantial barriers and cultural resistance to new scientific knowledge, and the role of moral and ethical beliefs" (National Research Council of the National Academies, 2012, p. 6). The National Academy of Engineering (2005) report adds that engineering students should be mindful "of the complexities associated with a global market and social context" (p. 10).

Teaching and discussing the social, political, and economics of one's discipline contributes to the fulfilment of WASC Senior College and University Commission accreditation requirement 2.2a. Specifically calling for the depth and breadth of a discipline to prepare the student for the workforce and life-long learning. It also expose students to be mindful of "diversity; ethical and civic responsibility; culture; social and political, and scientific and technical knowledge" (WASC Senior College and University Commission, 2013, p. 5).

Although there is an expressed need to teach and discuss social, political, and economic topics in a STEM curriculum, such an innovation is still in its infancy (Chong, Depew, Ngambeki, & Dark, 2013) as there are challenges. Challenges were identified while reviewing other data science programs from various universities, which were space and time to teach the topics. The purpose of this paper is to take a case study approach by introducing an example to integrate the social, political, and economic topics to a course called The Introduction to Data Science and Basic Statistical Programming from the Natural Science and Mathematics department at Chaminade University. The case study considered includes elements of a survey and focused course, in which the course embraces the survey topics of data science, social, political, and economics; and focused elements of programming and analyzing data in R, and interpreting and evaluating the results. The outline of this paper will discuss the process of developing a course to instruct the social, political, and economic topics by introducing Newell's (1994) instructional design framework; follow by a description on how the course implements the elements of the framework; linking the course to WASC Senor College and University Commission 2.2a accreditation requirement; and conclude with final thoughts and a next step.

The Introduction to Data Science and Basic Statistical Programming course used as the case study for this paper is one of six courses for a new data science certificate that will be offered by Chaminade University. The outcomes for this course were developed and guided by Bloom's Taxonomy of knowledge, comprehension, application, analysis, synthesis, and evaluation. The following are the outcomes of the introductory data science course:

- Identify and describe what is data science and why it is important socially, politically, and economically.
- Identify and describe the various types of databases, practices, and programming languages used.
- Identify and describe the process to collect, store, and curate datasets.
- Explain the process of data ingestion and wrangling.
- Understand the language basics of R.
- Apply R to perform various statistical analyses.
- Produce, synthesize, interpret, evaluate, and report results to various audiences.
- Discuss the social, political, and economic benefits and ramifications of the results.

Instructional design framework

Newell's (1994) instructional design framework was considered to design the data science course that integrates the social, political, and economic topics. The following are the elements of the framework:

- ***Choosing a topic and subtopics:*** *"It is the first step in designing a course. Should point back to the course desired learning outcomes and relate to the students in the course. A topic is a specific area, issue, theme, problem, region, time period, institution, figure, work, concept, or idea." A topic can have one or many subtopics or nested topics.*
- ***Identifying disciplines:*** *"Determine the appropriate disciplines that the course will use."*
- ***Developing the subtexts:*** *"The abstract issues of which the topic of the course is particularly embodied. Subtexts transpire what the course is about."*
- ***Structuring the course:*** *"The breadth, depth, and ordering of subtopics and subtexts that accumulates to a topic or subtopics."*
- ***Selecting the readings:*** *"Select readings that addresses students' limitations of the course, relates to students' interest, does not take up a lot of time, and reflects the different disciplinaries."*
- ***Designing the assignments:*** *"Assignments and activities should promote the desired educational outcomes."* (Newell, 1994, pp. 38–50)

Instructional design framework implemented to design data science course

This section presents details on how the Newell's (1994) instructional design framework is used to develop the introductory data science course. This section will first discuss the choosing of the topic and subtopics. Followed by identifying the disciplines. Next, present where and how the social, political, and economic topics are integrated through the subtexts and

structing of the course. This section will end on selecting the readings and designing the assignments.

Choosing a topic and subtopics

The topic of the course is the Introduction to Data Science and Basic Statistical Programming. The course includes several subtopics. The first is a survey of the components of data science (e.g., practices, databases, mindsets, languages, statistics, so forth). Another subtopic includes datasets through collection and handling (i.e., storing, curating, ingestion, and wrangling). The third subtopic is using R to analyze the datasets. The fourth subtopic is handling (i.e., produce, synthesize, interpret, evaluate, and report) the results of the analyses. The topic and subtopics were derived from a qualitative study that evaluated the data science programs from various universities; and large and small companies' responsibilities, qualifications, requirements, and skills.

Identifying disciplines

Since this course is open to the all majors who are interested in data science, the disciplines will draw from the expertise of the course instructor, Chaminade University degree programs, outcomes of the course, and available resources. Examples of disciplines and not an exhausted list could include historical and political studies, business administration, international studies, biology, environmental studies, biochemistry, or criminal justice. However, this course also presents other disciplines that are not offered by the University, but various faculty, staff, and University partners could be utilized for a class session or co-instructor of the course to provide further expertise.

Developing the subtexts

The subtexts on the topic of the Introduction to Data Science and Basic Statistical Programming, goes beyond the surveying of data science, the collection and handling of datasets, the programming, and the performance of analysis and results. The social environment introduced by Dark, Ngambeki, Depew, and Chong (2014) provides a way to think about subtexts for a course. According to Dark et al. (2014), a social environment is a context. Within this context includes three man-made systems where social, political, and economic opportunities and challenges emerges. An example, if the analysis found that there was an overfishing of a particular fish. The discussion can further delve into social (e.g., who gets affected), political (e.g., what are the current policies on overfishing and any suggestions for new policies), and economic (e.g., what are the changes in market prices) issues that can lead to further research.

Structuring the course

The structuring of the course is illustrated in Table 1. In Table 1, the sequence of subtopics is in the order of 1) surveying the components of data science, 2) datasets collection and handling (i.e., collect, store, curate, ingest, and wrangle), 3) R programming and analyses, and 4) handling the results from the analyses.

Table 1 Course structure

Topic	Subtopics	Subtexts discussions
Introduction to data science and basic statistical programming	1. Survey the components of data science.	Social, political, and economic challenges and opportunities.
	2. Dataset collection and handling	
	3. R programming and analysis	
	4. Handling the results from each analysis	Social, political, and economic challenges and opportunities.

The two areas the subtexts are included for this course would be the surveying the components of data science. The second area would be after each analysis (i.e., handling the results), during the interpreting and evaluating of the results as each type of analyses and related results may provide different social, political, and economic conversations.

Selecting the readings

The course has required readings that covers the data science subtopics and social environment systems subtexts. The readings are short, comprehensible, and will fill the gaps of student knowledge and skills. The accumulation of the readings shows a larger picture of the discipline to the students. Along with the readings, this course includes YouTube videos and websites to facilitate student learning and understanding of the topic, subtopics, and subtexts. This class has a semester long project and students will select the readings based on their diverse interests and disciplines.

Designing the assignments

The assignments considered three factors during the design, learner-centered design, project-oriented design, and constructive alignment. The goals for integrating these factors, students will co-create knowledge and skills in data science and the social environment, professors instructing a course becomes a guide, establishes a learner environment, and students become life-long learners. The assignments in this course illustrated in Table 2 includes two response papers early in the semester, eight programming and analyses assignments mid to end of semester, a semester long final project introduced the third week of semester, project proposal mid semester, and project update end of the semester. Both the analyses assignments and semester long project will use unclassified or publicly available datasets. Designing each assignment, the response papers, the programming and analysis assignments, project proposal, and project update all accumulates to the final project.

Table 2 Course assignments

Assignments	Early in the semester 1-4 weeks	Middle of the semester 5-10 weeks	End of the semester 11-16 weeks
Response papers (2x)	X		
Programming and analysis assignments (8x)		X	X
Project proposal		X	
Project update			X
Final project	X	X	X

Linking course to accreditation requirement 2.2a

A method to examine and satisfy the WASC Senior College and University Commission accreditation requirement 2.2a is by using guiding questions that could be organized in one or many social environment systems (i.e., social, political, and economic). The guiding questions can be used in the two areas mentioned, the structure of the course section and after obtaining the results from the various analyses. The guiding questions can be used during lectures, discussions, assignments, and the project. Below are potential guiding questions:

- Surveying data science
    - Why is data science important to <them (i.e., the student), their family, their community, their culture, their home (e.g., state or country)>?
    - What are the benefits and consequences of data science using Deborah Stone's theoretical social goal framework on equity, efficiency, security, and liberty?
- Interpretation and evaluation of results from the analyses
    - **Intertemporal resource transfer**: "What are the trade-offs of short-term impact and long-term consequences" (Jacobs, 2011, p. 3)? Are there any policies in place to prohibit short and longer-term consequences. If not, based on the results, what policy would you suggest to an authoritative figure (e.g., congress, manager, CEO).
    - **Pareto-efficient solution**: Does the <group analyzing> improve <situation> without harming or weakening others (Investopedia, 2018a)?
    - **Window of opportunity**: When would be the best time to promote the results to a <stakeholder> (Kingdon, 2011).
    - **Tragedy of the commons**: Does the demand of <resource> overwhelms the supply (Investopedia, 2018b)? Who benefits and who gets harmed?

Note: The symbol < > means that the instructor can customize the question.

Conclusion and next step

The National Research Council, the National Academies, and the WASC Senior College and University Commission calls for social topics to be integrated in STEM-oriented curriculum, in this case was an introductory data science course. This paper began by introducing a course design framework used to design the course. Next, the implementation of the course design framework discussed where and how to integrate the social, political, and economic topics. An approach was to introduce the topics as subtexts by connecting them with the subtopics of surveying data science, and at the ending of every analysis during the interpretation and evaluation of the results. This approach is not the only approach, Ngambeki et al. (2013) provides other methods to integrate the topics; for example, a short talk, module, or workshop. This paper is to introduce our approach in designing the data science course with the intent to include social, political, and economic topics. The next step will be to develop either a module or workshop (i.e., introduce some components of the course) as a pilot study based on the ASSURE instructional design framework as it has components of course execution and evaluation. The ASSURE framework is short for "analyze learner characteristics; state objectives; select, modify, or design materials; utilize materials; require learner response; and evaluation" (Kurt, 2018, para. 2). The module or workshop will be evaluated through a survey to collect students' self-efficacy, career development, and design of course and items covered.

Acknowledgements

This work was supported by the National Science Foundation grants of NSF I-USE (DUE1525884), NSF INCLUDES (HRD1744526), and NIH INBRE (P20 GM 103466) for the development of a culturally-responsive data science, analytics, and visualization undergraduate curriculum and summer immersion program at Chaminade University of Honolulu in partnership with Texas Advanced Computing Center (TACC). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the NSF, NIH, Chaminade University of Honolulu, and Texas Advanced Computing Center.

The author thanks Dr. Helen Turner (Chaminade University), Dr. Mark Speck (Chaminade University), the Data Science Program team, and the Division of Natural Sciences and Mathematics at Chaminade University of Honolulu for their insights to the development of this article.

References

Chong, R., Depew, D., Ngambeki, I., & Dark, M. (2013). Teaching social topics in engineering: The case of energy policy and social goals (pp. 1–16). Presented at the 120th ASEE Annual Conference and Exposition, Atlanta, GA: American Society for Engineering Education. Retrieved from http://www.asee.org/public/conferences/20/papers/5936/view

Dark, M., Ngambeki, I., Depew, D., & Chong, R. (2014). Social engagement by the engineer. In E. Coyle & R. A. Simmons (Eds.), *Understanding the global energy crisis*. West Lafayette, IN: Purdue University Press.

Grasso, D., & Martinelli, D. (2008). Holistic engineering: The dawn of a new era for the profession. In G. Madhavan, B. Oakley, & L. Kun (Eds.), *Career development in bioengineering and biotechnology* (pp. 303–307). New York, NY: Springer.

Investopedia. (2018a). Pareto efficiency. Retrieved July 30, 2018, from https://www.investopedia.com/terms/p/pareto-efficiency.asp

Investopedia. (2018b). Tragedy of the commons. Retrieved July 30, 2018, from https://www.investopedia.com/terms/t/tragedy-of-the-commons.asp

Jacobs, A. M. (2011). *Governing for the long term: Democracy and the politics of investment*. Cambridge, MA: Cambridge University Press.

Kingdon, J. W. (2011). *Agendas, alternatives, and public policies* (Second). Glenview, IL: Pearson Education Inc.

Kurt, S. (2018). ASSURE: Instructional design model. Retrieved July 30, 2018, from https://educationaltechnology.net/assure-instructional-design-model/

National Academy of Engineering. (2005). *Educating the engineer of 2020: Adapting engineering education to the new century*. Washington, DC: The National Academies Press. Retrieved from http://www.nap.edu/catalog/11338.html

National Research Council of the National Academies. (2012). *Using science as evidence in public policy*. Washington, DC: The National Academies Press.

Newell, W. H. (1994). Designing interdisciplinary courses. *New Directions for Teaching and Learning*, *1994*(58), 35–51. https://doi.org/10.1002/tl.37219945804

Ngambeki, I. B., Depew, D. R., Dark, M. J., & Chong, R. C. (2013). Approaches to integrating policy into engineering education (pp. 1–14). Presented at the 120th ASEE Annual Conference and Exposition, Atlanta, GA: American Society for Engineering Education.

WASC Senior College and University Commission. (2013). Handbook of accreditation 2013 revised quick reference guide. Retrieved from https://www.wscuc.org/resources/handbook-accreditation-2013