Experiential Learning in Data Science Through a Novel Client-Facing Consulting Course

Arko Barman
Data to Knowledge Lab
Rice University
Houston TX, USA
arko.barman@rice.edu

Su Chen
Data to Knowledge Lab
Rice University
Houston TX, USA
sc131@rice.edu

Andersen Chang
Dept. of Statistics
Rice University
Houston TX, USA
atc7@rice.edu

Genevera Allen
Data to Knowledge Lab
Rice University
Houston TX, USA
gallen@rice.edu

***Abstract*—This Innovative-Practice Full Paper presents the curriculum development and our experiences in offering a client-facing consulting course in data science. Data science education has seen rapid growth over the past decade. To provide students with hands-on opportunities to work with real data, many data science programs have advocated for and implemented experiential learning opportunities throughout the curriculum, which has been shown in a wide variety of literature to have many benefits. Most experiential learning opportunities in STEM programs are provided through capstone and engineering design courses; this is becoming increasingly the case in data science programs as well where several universities have developed data science capstone programs in which students work with clients on the client's real-world data sets. While client-sponsored capstone projects are an exemplar of experiential learning, they may pose major challenges to implement and can be particularly resource-intensive for institutions; this is especially the case in data science where the legalities of data sharing may come with additional hurdles. Because of this, we were motivated to develop a novel client-facing data science consulting course that provides a unique experiential learning scenario to both undergraduate and graduate students while requiring much fewer resources and legalities. In our novel data science consulting course, groups of students work directly with real clients in a consulting clinic setting to provide data science guidance and short-term help with data science challenges. Through this process, students learn about the diversity of real-world problems in data science, how to lead consultations with clients effectively as a team, how to frame data science challenges and research possible solutions, and how to communicate solutions to clients in reports and presentations. We leveraged best practices in consulting courses developed in business school settings to design our course. Additionally, the consulting course serves as a community service initiative whereby researchers, clinicians, non-profit and government workers, and industry professionals benefit from the advice and short-term help provided through consultation. In this paper, we report how our consulting course is set up, how clients from both within and outside the university can seek help at the consulting clinic, and how the structure of the course enables students to have firsthand experience working on many real-world data science problems with clients. Finally, we discuss how student performance is assessed in this course, the lessons learned from offering this course, and recommendations for other data science programs in universities that wish to design similar courses.**

***Index Terms*—data science education, consulting course, experiential learning**

I. INTRODUCTION

Data science is an interdisciplinary field of study that leverages methods and tools from computer science, machine learning, statistics, electrical engineering, and applied mathematics to derive meaningful insights, observations, and inferences from large amounts of possibly complex data. Along with rapid growth in available data, the need for data scientists both in industry and academia has grown rapidly over the past few years. To fulfill this demand for professional data scientists, several universities worldwide have started departments, centers, or institutes dedicated to data science and have started to offer both undergraduate majors and minors, as well as graduate degrees in data science. The need for experiential learning in data science education has been widely recognized by instructors to be instrumental in providing valuable hands-on experience to students [1]–[5]. Most data science programs offer capstone courses to students, where they work on a data science problem over one to two semesters [1], [4]–[6].

Degree programs in data science have grown exponentially in recent years along with an increased interest in data science education in general. The design of new curricula for data science programs have been studied and reported in detail along with several recommendations for curriculum development [1]–[3], [5], [7], [8]. Several academicians have stressed the need to emphasize experiential learning in data science curriculum and have shown improved learning outcomes when opportunities for experiential learning are provided in the curriculum [2], [9], [10]. Students in data science are often engaged in experiential learning through solving real-world problems and by working on datasets from the real world. One of the ways to encourage students to employ the skills they have acquired over a series of courses is a capstone program where the students have the opportunity to work with real-world data [5], [8]. It has been shown that designing client-facing courses for solving real-world problems often helps in motivating students and keeping them engaged through more meaningful and realistic project objectives [11]–[14].

While a capstone course provides an excellent opportunity for students for an in-depth experiential learning experience, exposure to the breadth of data science problems and applications in the real world is elusive in this scenario. In this paper, we discuss the development and implementation of a novel experiential data science course that enables students to be exposed to a wide variety of problems related to data science over the course of a semester. We introduce a data science consulting course (stylized as the "Data Science Consulting Clinic" for clients), where enrolled students have the unique opportunity to consult with clients regarding their data science problems, hence creating a scenario where they gain exposure to a plethora of data science problems encountered in real-life applications. Besides experiential learning, the students are also exposed to other successful active learning paradigms, such as collaborative learning. In this paper, we describe our experiences in the design and development of this course using evidence-based experiential learning methodologies, and also discuss results from student feedback, lessons learned, and recommendations for other faculty members in the computer science and data science communities who might consider developing a similar course.

While some consulting courses are available for experiential learning for students in business and management [15], [16], operations research [17], and statistics [18], [19], data science curricula in universities rarely include a consulting course. The need for a consulting course in data science is only recently being realized by academicians and instructors [20], [21]. However, the structure of our data science consulting course differs considerably from the ones in other domains. For example, in management, consulting courses involve one of three methods [22]: (i) a one-day on-campus workshop followed by field study; (ii) business process consulting and simulation; and (iii) a strategy consulting workshop. On the other hand, consulting courses in statistics focus on completing entire research projects from start to finish, including the implementation of methods and the analysis of data, and thus involves solving larger but fewer problems over the course of the semester [19], [23]. In contrast, our data science consulting course focuses on hosting a consulting session with a client followed by a report with recommended solutions submitted to the client with a turnaround time of one week only.

Educational literature on data science experiential learning through a consulting course is very sparse. [24], [25] described statistical consulting courses for undergraduates at several institutions. These structured courses with a well-defined curriculum often rely heavily on "mock consulting" with internal clients or canned projects, rather than expose students to live clients and real-world challenges. In our consulting course, students work with real clients from day one, receiving constant feedback and mentoring from faculty to improve both their technical and soft skills. We document that such an approach, while resource-intensive, is proven to be extremely valuable for students with diverse backgrounds. While most statistical consulting programs or consulting centers, such as [26], [27], focus on project-based experience where students work on the clients' problem and implement solutions, [28] discussed the pros and cons of using two modes: clients with semester-long projects and drop-in clients with only short-term projects. By only working with drop-in clients who signed up for free consulting service, we expose students to a diverse client base and a broad range of data science problems. In addition, students are required to make recommendations but do not necessarily implement the solution. This allows students to focus on problem formulation and brainstorming, which differs significantly from their traditional learning experiences. [14] advocates the pedagogy that leverages boundary theory: position students to be at the boundary between real clients with real-world data science challenges and the academic environment. Our course provides a platform that encourages students to cross the boundaries and work collaboratively in interdisciplinary teams. In summary, we add to the chorus in favor of adding a consulting course to data science experiential learning and contribute to the literature by describing the novelty of our course design and reporting assessment of the course and feedback from students and clients.

II. Design & Development of the Consulting Course

A. Goals & Objectives

The primary goal of designing the consulting course was to provide experiential learning opportunities in data science to both undergraduate and graduate students. The objectives for developing this course are discussed here.

First, while students can gain an in-depth experiential learning experience through capstone courses in data science that have become common in recent years [1], [4], opportunities for expanding their breadth of experience related to data science problems in the real world are limited. In particular, the only opportunity for students to gain a breadth of experience in capstone courses is through presentations, seminars, and posters of other teams in the capstone program. However, attending presentations and seminars is inadequate for providing hands-on experience since the intricacies of handling and analyzing real-world data firsthand cannot be replicated by attending presentation sessions. Moreover, such presentations present both the problem and the solution, and, thus, do not encourage problem-solving or brainstorming for students in other capstone teams.

Our other major objective in designing the consulting course was to expose data science students to a variety of smaller or less complex data science problems that do not necessarily warrant a semester-long capstone project to be tackled. There are several such data science problems, e.g., some researchers might have collected data for certain experiments and are looking to select or design a suitable statistical hypothesis test to make conclusions. Such less complex problems allow students to work in groups to solve a data science problem for a short span of time while creating a client-facing scenario similar to client-facing capstone projects.

A secondary objective for developing the consulting course has been to encourage collaborative learning by forming teams of students to work on solving problems. This design is inspired by evidence showing the benefits of collaborative learning in client-facing projects for solving real-world problems [29]–[32]. Further, teams of students are formed so that each team comprises students from different levels of study, viz. graduate and undergraduate, as well as from different backgrounds, thus enabling teams to be interdisciplinary. The creation of interdisciplinary teams not only leads to improved learning outcomes [33] but also facilitates collaborative learning through the jigsaw approach [34].

A final objective of this course is to provide a unique service to the community of researchers, faculty members, clinicians, industry professionals, and non-profit groups through recommendations provided by the consulting to solve their data science problems. During the summer and in semesters when this course is not offered, the instructors continue to offer data science consulting sessions as a community service initiative.

The learning outcomes of the course are that the students completing this course should be able to (i) listen and ask questions of clients to ascertain the data science challenges, (ii) formulate challenges into coherent data science problems, (iii) provide guidance or simple solutions on the identified problems, and (iv) communicate with clients orally and in reports to help them understand the problems, solutions, and how to interpret results.

B. Curriculum

Despite the popularity of consulting courses in other fields of study [35], there is a dearth of literature outlining the design and development of a data science consulting course, details of its curriculum, and key reflections from offering such a course over several semesters. This course is set up as a lecture/lab course. The lectures include both lectures from instructors (see Sections II-B2 and II-B4) as well as student presentations followed by discussions on client problems from the clinic in that week (see Section II-C1). Here, we discuss in detail the curriculum design of our data science consulting course.

1) The Application Process: The consulting course has an application process that helps the instructors gauge the knowledge and exposure of the students to various areas in data science and soft skills that they might require during consulting sessions with clients. These areas include - (i) major of study; (ii) prior experience of working in teams; (iii) statistical analysis and modeling; (iv) database systems; (v) machine learning; (vi) data wrangling; (vii) signal processing; (viii) applied data science; and (ix) coding skills in R and Python. This questionnaire is designed to enable the instructors to understand the unique background of each student to determine the best students while forming teams to solve a particular client problem. Occasionally, students with no background in machine learning and/or statistics are asked to drop the course.

2) Introductory Lectures/Sessions: The first two weeks of the course involve explaining the structure of the course and its logistics to enrolled students, as well as making them aware of the best practices in consulting. Since it is a non-traditional course with its primary focus being to solve problems brought by clients in an impromptu fashion, the goal in the initial weeks is to make the students acquainted with the process of consulting through lectures on how to understand the client problem, how to determine the objectives of the client, and how to convert the client objectives into data science objectives. Additionally, lectures and resources on the best practices of how to conduct research or brainstorm on new topics, how to provide recommendations to clients, how to effectively conduct client meetings, and how to write professional reports for clients are provided to the students. Any clients coming to the consulting sessions for the initial two weeks are handled by the instructors while the students are asked to observe and take notes. These sessions offer a practical demonstration vis-à-vis the introductory lectures on how to conduct consulting sessions and the best practices in consulting.

3) Regular Consulting Sessions: Following the initial two weeks, students are assigned for hands-on consulting with clients. For every client that comes to the clinic, the instructors create a team of two to four students, usually with the requisite background needed to advise the client on their problem. Every consulting session is supervised by at least one of the instructors to ensure that both client and the students have the best possible experience from the consulting session. While the consulting sessions are led by the students themselves, the instructor(s) may intervene from time to time to guide the consulting session toward a fruitful direction.

4) Lectures on Data Science Topics: Since students from any background can enroll in the data science consulting course, not all students possess an adequate grasp of all topics that might help them conduct consulting sessions successfully. In addition, some topics that students might not have encountered previously in their courses, such as mixed effects models, occur frequently in client problems. To help students fill in these potential gaps in knowledge, the instructors deliver lectures on relevant data science topics from time to time. These lectures cover a wide variety of topics ranging from statistical models, e.g., survival analysis, to deep learning models, e.g., convolutional neural networks.

C. Student Deliverables & Assessment

Students are required to submit certain deliverables to the instructors and others to the clients they have consulted for. Here, we describe the student deliverables in detail.

1) Client Reports & Class Presentations: For every client attending a consulting session, the student team consulting for them is required to submit a draft report to the instructors for their review. After the instructors have provided their feedback and corrections to the draft, the student team sends the final consulting report to the client. A template is provided to the students for preparing the report so that all reports have the same structure. The preparation of client reports is an exercise designed to help the students develop and improve their written communication skills. Further, the student team also requests the client to complete a survey detailing their experience in the consulting session and provide their feedback for improvement of both the curriculum and the students.

Recommendations on the best ways to implement the solutions are also provided to clients. For example, application programming interfaces (APIs), packages, and libraries in R or Python are recommended for every solution recommended by students, if applicable. It must be noted here that this consulting course is designed so that only recommendations for solving problems are provided to clients. The implementation of these recommendations is not performed by students, except possibly for a final project that we discuss in detail later. This is another aspect of this course that makes it distinct from data science capstone courses, where student teams are usually required to develop software for solving a problem.

In addition, every student team presents a summary of the consulting session during the lecture portion of the class. This presentation provides a means for other students in the class to acquaint themselves with the myriad of data science problems they have not consulted for. As a result, every student in the consulting course becomes familiar with every consulting client and every problem that is being solved as a part of the consulting course. Moreover, this also creates the opportunity for students to improve their oral communication skills. The students in the audience are encouraged to engage in discussion with the presenting team and the instructors. Often, in our experience, these presentations lead to interesting discussions not only about the nature of the specific problem and the recommended solutions but also about broad application areas of data science that the students have not had prior exposure to, such as natural language processing and computer vision.

2) Student Reflections: The consulting course is designed to provide students an exposure to a wide variety of datasets and data science problems encountered in real life. We envisage that the students can learn through an experiential learning setup that enables them to apply the knowledge that they have acquired to the real world. Keeping this in mind, we require students to submit a midterm reflection and a final reflection where they discuss the problems that they have encountered in the course of the semester. Students are encouraged to express their thoughts on which problems they found particularly challenging, which new areas of data science and its applications they were exposed to, and what topics they had to research for making recommendations to clients for solving their problems. This is an opportunity for students to assess their capabilities and growth, as well as another means of helping students improve their written communication skills. In particularly busy semesters with a large number of clients, the instructors sometimes decide to forego the midterm reflection to ease the workload on the students.

3) Final Project: The consulting course requires students to undertake a final project for the last three weeks of the semester. For this project, students can either choose to work on a consulting problem with a client from earlier in the semester or a project of their choice related to a previous consultation with the approval of an instructor. If a student chooses to work on a consulting problem, they can reach out and work with a previous client to implement the recommendations that were provided through the consulting session and the report. If a student chooses to work on a publicly available dataset, they must work with an instructor to define a suitable problem related to a previous consultation that can be solved or implemented within the stipulated time of the project. Usually, these topics for projects are chosen by students and are often interesting applications of data science, such as financial analyses of stocks and analyses of adoption rates in animal shelters.

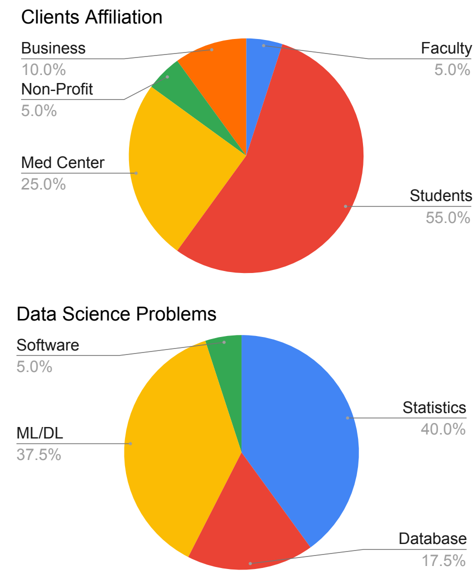

Fig. 1. Percentage of clients affiliation (top) and data science problems consulted (bottom) from Fall 2020 to Spring 2021 (n = 40).

4) Student Assessment: While every deliverable for the course has a component in student assessment and the final grading, particular attention is also paid to participation in the consulting sessions, individual contributions to the team, and participation during the lecture or discussion sessions. In addition, the instructors also focus on communication skills, both written and oral. A grading rubric has been developed which balances the different components of assessment. The assessment and grading focus on the following main aspects: (i) Attendance & participation; (ii) Written consulting reports; (iii) Oral presentations; and (iv) Final project & reflections.

D. Clients

The consulting clinic is open to the university community and beyond, including a Medical Center, non-profit organizations, and for-profit businesses. Due to the pandemic, the format of the clinic has changed from walk-in to Zoom sessions with appointments. The diverse client base has significantly broadened the type of data science problems they have brought in. Here we summarize the type of data science problems into four broad categories:

- Traditional statistical analysis: this category includes experimental design, sample size calculation, hypothesis testing, bio-statistics, survival analysis, and similar topics.
- Machine learning and deep learning: this category includes predictive models for regression and classification, computer vision, signal processing, anomaly detection, natural language processing, and related problems.

Client type	Database	ML & DL	Software	Statistics	Total
For-profit business	1	2		1	4
Non-profit organization		1		1	2
University faculty				2	2
University grad student	2	10	2	5	19
University undergrad	2	1			3
Medical center	2	1		7	10
Grand total	7	15	2	16	40

TABLE I
A PIVOT TABLE SHOWING THE CATEGORIZATION OF TYPES OF PROBLEMS PRESENTED AT THE CONSULTATION VIS-À-VIS TYPES OF CLIENTS. (ML STANDS FOR MACHINE LEARNING, DL STANDS FOR DEEP LEARNING)

- Database-related problems: this category includes data acquisition, data engineering, database design, and management.
- Software-related problems: this category usually refers to specific questions about software, e.g., how to implement certain models in R or Python.

Figure 1 shows the percentage of clients' affiliations and data science problems categorized as above during the semester of Fall 2020 and Spring 2021. A more detailed categorization of problems vis-à-vis the type of clients is shown in Table I. Data reflects the most recent academic year, not the entire history of the consulting clinic.

III. Results & Lessons Learned

We designed and offered this course at a small private university with a total enrolment of fewer than 8,000 students. In addition, this course is offered as a fully free elective with no majors and minors having this course as an elective in their prescribed curriculum. Based on this context, we present our results both from the perspective of students and clients in our consulting course.

A. Student Self-Assessment

One of the key components of midterm and final reflection assignments is to give students an opportunity to appraise their own experience within the course. In particular, as part of the midterm reflection, we ask students to describe what skills they would like to improve upon during the second half of the course. Then, in the final reflection, we ask students to assess their improvement over the course of the semester.

In general, we observe two main areas in which students feel that they struggle the most during the midterm reflection. Most commonly, students comment that they feel the need to improve their communication skills, especially with regard to asking questions or expressing ideas to the client during the consulting session. For example, one student commented:

> "My biggest weakness in the consultations is that when I am stuck in a problem, I haven't been very good at reframing the question to be able to think about it in another way. [...]"

Another aspect of the consulting course that students mention that they can improve is the technical knowledge of statistical methods or terminology that commonly occur during client problems. In particular, students may know certain facets of data science, but not necessarily others that may come up in certain consultations:

> "[...] I need to better my knowledge in experimental design and inferential statistics. Before coming to this class, I mostly focused on learning different machine algorithms for prediction tasks, as I believed that most of my consultations would require extensive knowledge in data mining and learning. However, this has not been the case during my time at the clinic, as most of my clients need help in either data visualization or performing inference on their results."

In their final reflections, students almost unanimously reported that they had noticed improvements in both their communication and technical skills over the semester. For example, in the comments left in the final reflections by the same two students quoted above, we see that they felt that their capabilities had dramatically improved since their midterm reflection:

> "[...] I learned a great deal of alternative methods for analysis, such as mixed effect models, convex clustering, and t-SNE. In addition, I learned a great deal of new methods for inference, mostly through my work with clients."

> "I improved over the semester in my confidence, organization, and the ways in which I attempted to help clients. [...] my confidence going into consultations has improved as I felt more comfortable interacting with clients and communicating to them what my opinions on the matter are. [...] the overall structure of going into the consultations were a lot better as we gained experience."

B. Student Demographics & Feedback

From Spring 2018 to Spring 2021, the consulting course had a total of 58 students. In terms of education level, 34 of these were undergraduate students, while 11 were Ph.D. and 13 were master's students. In terms of background, 29 of the students came from statistics, 10 came from computer science, 10 from other engineering programs, and the rest from a variety of majors including bioscience, economics, and business.

Student feedback for the course is provided to instructors through anonymous final course evaluations. These gathered for the course via university surveys at the end of each semester. For these evaluations, students rate the course on several metrics, including overall course quality and how much the course challenged the student to extend or develop their capabilities. We have aggregated these ratings across all semesters of the course, in which 53 out of the 58 students responded to the survey. We show an example of results in Figure 2; generally, we found that students responded

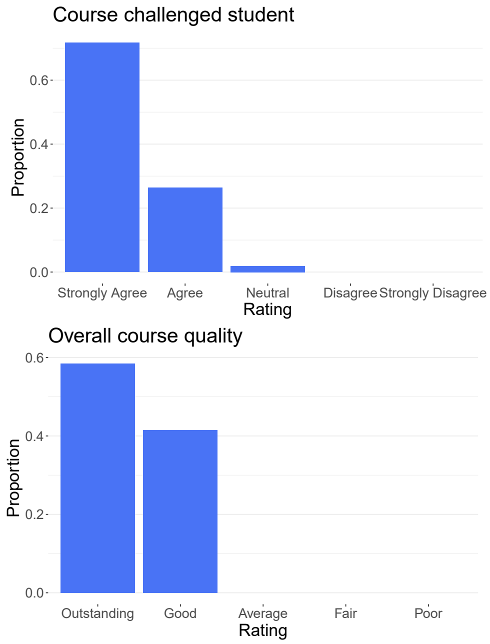

Fig. 2. Aggregated student ratings of course challenge (top) and quality (bottom) from Spring 2018 to Spring 2021 (n = 53).

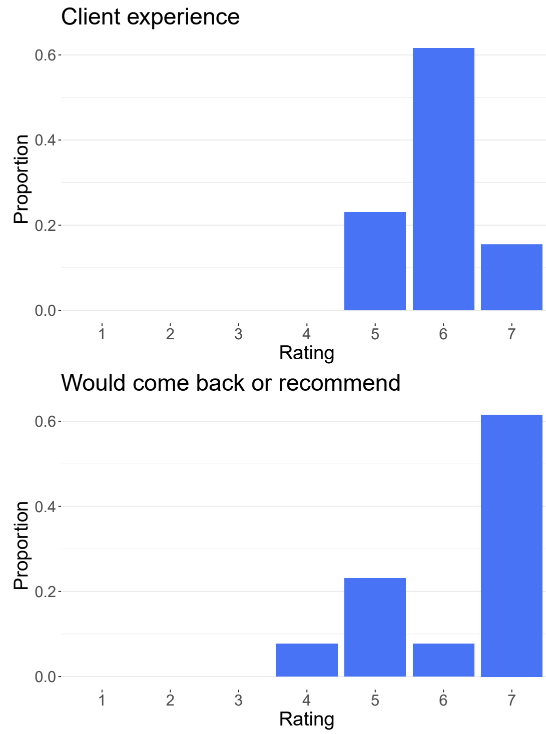

Fig. 3. Aggregated clients ratings of "overall experience": rating 1=unhelpful, 7=extremely helpful! (top) and "would come back or recommend": rating 1=strongly disagree, 7=strongly agree (bottom) from Fall 2019 to Spring 2021 (n = 13).

positively to the university course evaluations with regard to the survey metrics.

Students are also given the opportunity to leave anonymous comments describing their experience with the course, written toward an audience of future students who may consider taking the course. Overall, students most often tended to praise the experiential learning component of the course, while also occasionally suggesting improvement in course organization. Some examples of comments left by students:

> "This will be different from the other stat electives as you will actually put your stat skills to use in the real world. As you interact with more clients, you will also gain more experience in translating a practical problem into a statistical problem, and how to communicate that to people without stat background."
>
> "I thought this class was phenomenal. Getting the opportunity to interact with clients and their real-world problems was great for getting consulting experience, and applying what I had learned to industry problems. I felt like I was constantly learning new things and expanding my knowledge."

C. Feedback from Clients

Clients' experiences and feedback are collected through a follow-up survey using Google form. The survey link is sent to the clients together with the consulting reports via email and includes the clients' self-reported ratings on the consulting experiences. Due to the voluntary nature of this survey, we have collected limited data (n=13) and the results might reflect potential selection bias:

- 77% of clients rated the overall experience 6 or above out of a 1 to 7 Likert scale, with 1 being unhelpful and 7 being extremely helpful. None rated below 5.
- 92% of clients would come back to the consulting clinic or recommend it to their colleagues based on a 1 to 7 Likert scale.

Figure 3 shows the distribution of previous data on client satisfaction ratings. Client feedback is not taken into consideration for the assessment of students.

D. Evolution based on student and client feedback

Students are encouraged to provide comments and suggestions directly to the instructors to help improve the course and accommodate specific needs. This feedback has helped the instructors improve several facets of the consulting clinic and the course:

- For many consultations, students help provide clients with additional online resources or references to help them implement the suggestions provided by the consulting teams. To help with this, course instructors compiled a developing list of standard resources for students to use for consultations.
- In the midterm reflections, some students have expressed interest in participating in consultations with clients from specific fields, such as finance or biomedical. Course instructors attempt to assign students to teams based on these requests when possible.

In addition, one survey question asks clients to provide any feedback on how we can improve the Consulting Clinic and any additional comments in free text. We have implemented the following changes according to the clients' feedback:

- Before the outbreak of the COVID-19 pandemic, the consulting clinic only served clients on a walk-in basis during the consulting session each week, i.e., no appointment required. After moving consulting sessions to online mode, the consulting clinic requires clients to fill out an appointment request with a brief description of their problem. This provided the students an opportunity to get an overview of the client's project and familiarize them with the specific field before the consulting takes place.
- We now provide clients with some instruction to prepare for the consulting when they sign up for an appointment, e.g., resources to bring to the consulting session that might be helpful, e.g., mock data and example analyses from literature.
- Student teams now provide more explanation on their recommendations for clients who have follow-up questions via email and clients are encouraged to book a follow-up consulting appointment if needed.

E. Case Studies

During the regular data science consulting sessions, students can face a variety of technical as well as communication and teamwork challenges. One particularly difficult consultation involved a client who had data on patients with a specific heart condition that had undergone a sequence of surgeries. The data included demographic information, as well as levels of various chemicals in the blood of each patient that were measured at irregular intervals; the client wanted to determine if there was any relationship between these variables and patient outcomes. During the consultation, the student team had difficulty grasping the potential biological relationships among the different measured variables and with the patient outcomes due to a lack of domain knowledge. Additionally, the team struggled to understand the non-standardized measurement times and how to model this. Because of these difficulties, the team members reached different understandings of the research problem and struggled with translating the client problem to a data science problem. Eventually, the session required instructor intervention as well as extra time spent after the consultation to reach a consensus on how to frame the problem and find solutions for the client.

In another consulting session, a client produced sun spectroscopy measurement data in the form of electromagnetic wave readings. The client wanted to reconstruct missing observations from the current data as well as to predict wave amplitudes in future spectroscopy observations. The client's problem involved the usage of data imputation techniques to replace the missing observations in the current data and supervised learning models to forecast future observations. The student team for this consultation had little familiarity with data imputation methods and methods for modeling spectroscopy or time-dependent data. Through a follow-up lecture by an instructor and research done after the consulting session, the students were able to recommend several clustering and regression-based imputation methods for data imputation. The team also recommended neural network models for classifying spectroscopy waves and predicting occurrences of spectroscopy patterns on future days.

Both of these case studies exemplify common challenges students face in the consulting clinic: (i) lack of background domain knowledge to understand the client problem or difficulty translating client terminology, (ii) challenges with framing client problems as data science problems, especially if the problem is completely unfamiliar, and (iii) limited technical expertise and experience in the data science tools necessary to solve the client problem.

F. Lessons Learned

We present our experiences and lessons learned from having offered this course from Spring 2018 to Spring 2021.

1) Improving Communication Skills with Clients: We noted that in the initial weeks of the semester, communicating with clients was a challenge for students due to a lack of prior experience. However, as the semester progresses, we observe a marked improvement in the communication skills of the students. Following the initial weeks of the semester, students understand the need to ask clarifying questions to the client, how to frame them, and how to translate the problem in the original domain to a data science question.

2) The Jigsaw Model: As the enrolled students hail from a variety of backgrounds, students learn from the expertise and knowledge of the other students who are assigned to their team for a particular client. In addition, the breadth of skill sets and prior experience across a team often helps in the "cross-pollination" of ideas and results in a better experience for the client as well.

3) Background of the Instructors: Due to the interdisciplinary nature of data science, we found it particularly helpful if the instructors have varied backgrounds and areas of expertise. The course was offered in multiple semesters with 1, 2, and 3 instructors. Having instructors with diverse backgrounds, including statistics, machine learning, deep learning, computer vision, natural language processing, time series analysis, and signal processing was found to be very beneficial for the consulting sessions.

4) Intervention of Instructors during Consulting Sessions: Occasionally, students are unable to tackle the problem posed by a client at a consulting session. It is crucial for the instructors to closely supervise the consulting sessions in such a way that they step in whenever they deem that the session is not moving towards a fruitful solution. These cases occur throughout the semester but are more prone to occur in the earlier weeks when the students have not yet had sufficient experience in consulting. It is often a fine balance on how and when the instructors should intervene in the consulting sessions.

5) Written Reports for Clients: Written reports submitted by students to clients (after being reviewed by the instructors) are widely appreciated by the clients. Through a written record of recommendations and solutions along with suitable references, clients have a document that outlines their problem and the recommended solutions that they can later implement.

6) Advertising the Consulting Clinic: One of the most crucial aspects of developing this course was to ensure a steady stream of clients throughout the semester to provide the students with a rich experience of a wide variety of problems. It is imperative that the consulting clinic be widely publicized among researchers and professionals within the university, among the industry, among other universities, hospitals, research centers, and non-profit organizations. To this end, we used help from our administrators in publicizing the consulting clinic as widely as possible. The students were also asked to help in advertising the clinic through social media, flyers, and mailing lists.

7) Workload for Students: We ensure that the workload is balanced among all the students of the class by meticulously keeping track of the number of consulting sessions each student is involved in, the difficulty of these sessions, and the contributions of the students in these sessions. In general, we limit every student to only one consulting session a week. However, if there are several clients, it is sometimes necessary to assign a student to two different consulting teams (with non-overlapping time slots) in one week. These occurrences are compensated and balanced by ensuring a lower workload for that particular student in the following week and a similar number of clients for all students throughout the semester.

8) Waiver of Liability: We prepared a waiver of liability and a consent for consultation form for clients both within and outside the university. These forms need to be signed by the clients before their consulting session. In particular, these forms allow us to work with outside clients without any liability for our consulting.

IV. Discussion

In this paper, we have discussed the design and development of a curriculum for a novel data science consulting course. A description or experience report of such a course for students studying data science in universities has not been reported yet, to the best of our knowledge. This course offers students the opportunity to be acquainted with the breadth of real-world data science problems encountered by researchers, faculty members, industry professionals, non-profit groups, clinicians, and other experts. While the primary focus of the course is to provide experiential learning to data science students, this course additionally provides an invaluable and unique service to the community. Moreover, implementation of a data science consulting course requires much fewer resources and legalities compared to other experiential learning courses, such as a capstone course, which involves complex logistics, such as access to confidential data, non-disclosure agreements, and intellectual property rights.

The unique challenges of a data science consulting course vis-à-vis other consulting courses in business and statistics are manifold. The diversity of the problems, the nature of our consulting clinic where we provide only recommendations and no implementations or simulations, and a turnaround time of one week to submit a report to the client outlining recommendations for solving their problem are some of the major challenges in this course.

In our experience, students have found the course to be pleasantly challenging and enriching to their knowledge and expertise in data science. On a similar note, most clients have found the consulting sessions to help solve their problems. We report and discuss our results based on the feedback from a few semesters. However, we will continue to collect feedback from students and clients in the future to keep assessing the usefulness of the course and evaluating how the course can be updated and refined in future semesters. In addition, we are working to scale up the course both in terms of the number of students and clients and use more faculty resources in expanding the reach of the course. In the future, we also plan to work towards including this course as an elective in relevant majors and minors offered by the university.

The course fits into the data science curriculum both at undergraduate and graduate levels, where the students enrolling must have prior knowledge of statistics and/or machine learning, along with other optional prior knowledge of visualization and database systems. Students wishing to understand the breadth of data science and its applications in different domains would be the ones to benefit most from such a course. While this course was offered with small enrollment, it can be scaled up given there is enough outreach to attract clients and enough instructors with diverse backgrounds to advise students.

In conclusion, we hope to encourage other universities and programs to introduce data science consulting courses for the benefit of students and clients. We would like to emphasize the interdisciplinary and experiential nature of the course as two of its primary benefits. We recommend the participation of faculty members from diverse backgrounds and a wide range of expertise for the best learning outcomes. We also recommend that before starting this course, the instructors and the program administrators create a network of universities, departments, centers, institutes, companies, and non-profit organizations where they can publicize and advertise the clinic to ensure that students have an adequate number of clients throughout the duration of the course.

References

[1] P. Anderson, J. Bowring, R. McCauley, G. Pothering, and C. Starr, "An undergraduate degree in data science: curriculum and a decade of implementation experience," in *Proceedings of the 45th ACM technical symposium on Computer science education*, ser. SIGCSE '14. New York, NY, USA: Association for Computing Machinery, Mar. 2014, pp. 145–150. [Online]. Available: https://doi.org/10.1145/2538862.2538936

[2] R. D. De Veaux, M. Agarwal, and et al., "Curriculum Guidelines for Undergraduate Programs in Data Science," *Annual Review of Statistics and Its Application*, vol. 4, no. 1, pp. 15–30, 2017. [Online]. Available: https://doi.org/10.1146/annurev-statistics-060116-053930

[3] N. Engineering, and Medicine, *Data Science for Undergraduates: Opportunities and Options*. Washington, DC: The National Academies Press, 2018. [Online]. Available: https://www.nap.edu/catalog/25104/data-science-for-undergraduates-opportunities-and-options

[4] G. I. Allen, *Experiential Learning in Data Science: Developing an Interdisciplinary, Client-Sponsored Capstone Program*. New York, NY, USA: Association for Computing Machinery, 2021, p. 516–522. [Online]. Available: https://doi.org/10.1145/3408877.3432536

[5] S. Rosenthal and T. R. Chung, "A Data Science Major: Building Skills and Confidence," in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, ser. SIGCSE '20. New York, NY, USA: Association for Computing Machinery, Feb. 2020, pp. 178–184. [Online]. Available: https://doi.org/10.1145/3328778.3366791

[6] R. A. Irizarry, "The Role of Academia in Data Science Education," *Harvard Data Science Review*, vol. 2, no. 1, Jan. 2020. [Online]. Available: https://hdsr.mitpress.mit.edu/pub/gg6swfqh/release/1

[7] J. Hardin, R. Hoerl, N. J. Horton, D. Nolan, B. Baumer, O. Hall-Holt, P. Murrell, R. Peng, P. Roback, D. T. Lang, and M. D. Ward, "Data Science in Statistics Curricula: Preparing Students to "Think with Data"," *The American Statistician*, vol. 69, no. 4, pp. 343–353, Oct. 2015. [Online]. Available: https://doi.org/10.1080/00031305.2015.1077729

[8] B. Ramamurthy, "A Practical and Sustainable Model for Learning and Teaching Data Science," in *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, ser. SIGCSE '16. New York, NY, USA: Association for Computing Machinery, Feb. 2016, pp. 169–174. [Online]. Available: https://doi.org/10.1145/2839509.2844603

[9] J. Silva, R. Portillo, and et al., "Data Sciences and Teaching Methods—Learning," in *International Conference on Communication, Computing and Electronics Systems: Proceedings of ICCCES 2019*, V. Bindhu, J. Chen, and J. M. R. S. Tavares, Eds. Singapore: Springer, 2020, pp. 335–342.

[10] E. Serrano, M. Molina, D. Manrique, and L. Baumela, "Experiential Learning in Data Science: From the Dataset Repository to the Platform of Experiences." in *13th International Conference on Intelligent Environments - IE'17*, Aug. 2017.

[11] B. Bruegge, S. Krusche, and L. Alperowitz, "Software Engineering Project Courses with Industrial Clients," *ACM Transactions on Computing Education*, vol. 15, no. 4, pp. 17:1–17:31, Dec. 2015. [Online]. Available: https://doi.org/10.1145/2732155

[12] L. Johns-Boast and S. Flint, "Simulating industry: An innovative software engineering capstone design course," in *2013 IEEE Frontiers in Education Conference (FIE)*, Oct. 2013, pp. 1782–1788.

[13] M. E. Beier, M. H. Kim, and et al., "The effect of authentic project-based learning on attitudes and career aspirations in STEM," *Journal of Research in Science Teaching*, vol. 56, no. 1, pp. 3–23, 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/tea.21465

[14] J. Saltz and R. Heckman, "Big Data science education: A case study of a project-focused introductory course," *Themes in Science and Technology Education*, vol. 8, no. 2, pp. 85–94, Feb. 2016. [Online]. Available: https://www.learntechlib.org/p/171521/

[15] R. LeMaire, K. Fisher, and E. F. Watson III, "Delivering an mba management consulting experience," *Decision Sciences Journal of Innovative Education*, vol. 15, no. 2, pp. 138–160, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/dsji.12124

[16] A. Richter and S. L. Schmidt, "The effectiveness of university-level management consulting courses," *Journal of Management Education*, vol. 32, no. 1, pp. 84–99, 2008. [Online]. Available: https://doi.org/10.1177/1052562907300809

[17] K. C. Heriot, R. Cook, R. C. Jones, and L. Simpson, "The use of student consulting projects as an active learning pedagogy: A case study in a production/operations management course," *Decision Sciences Journal of Innovative Education*, vol. 6, no. 2, pp. 463–481, 2008. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-4609.2008.00186.x

[18] A. P. Sima, V. A. Rodriguez, K. E. Bradbrook, B. S. DiPace, and V. Okhomina, "Incorporating professional recommendations into a graduate-level statistical consulting laboratory: A case study," *Journal of Clinical and Translational Science*, vol. 5, no. 1, p. e62, 2021.

[19] D. L. Hydorn, "Engaging Students in the Practice of Statistics Through Undergraduate Research," *Journal of Statistics Education*, vol. 26, no. 2, pp. 87–91, May 2018. [Online]. Available: https://doi.org/10.1080/10691898.2018.1475204

[20] I. Bile Hassan, T. Ghanem, D. Jacobson, S. Jin, K. Johnson, D. Sulieman, and W. Wei, *Data Science Curriculum Design: A Case Study*. New York, NY, USA: Association for Computing Machinery, 2021, p. 529–534. [Online]. Available: https://doi.org/10.1145/3408877.3432443

[21] J. Lasser, D. Manik, A. Silbersdorff, B. Säfken, and T. Kneib, "Introductory data science across disciplines, using Python, case studies, and industry consulting projects," *Teaching Statistics*, vol. 43, no. S1, pp. S190–S200, 2021. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1111/test.12243

[22] S. L. Schmidt and A. Richter, "Course formats for teaching management consulting," *Journal of education for business*, vol. 82, no. 1, pp. 56–62, 2006.

[23] B. J. Smucker and A. J. Bailer, "Beyond Normal: Preparing Undergraduates for the Work Force in a Statistical Consulting Capstone," *The American Statistician*, vol. 69, no. 4, pp. 300–306, Oct. 2015. [Online]. Available: https://doi.org/10.1080/00031305.2015.1077731

[24] K. Boomer, N. Rogness, and B. Jersky, "Statistical Consulting Courses for Undergraduates: Fortune or Folly?" *Journal of Statistics Education*, vol. 15, no. 3, p. null, Nov. 2007. [Online]. Available: https://doi.org/10.1080/10691898.2007.11889542

[25] R. H. Taplin, "Teaching Statistical Consulting Before Statistical Methodology," *Australian & New Zealand Journal of Statistics*, vol. 45, no. 2, pp. 141–152, 2003. [Online]. Available: http://onlinelibrary.wiley.com/doi/abs/10.1111/1467-842X.00270

[26] H.-J. Kim, K. S. Alberts, and S. Thatcher, "TEACHING UNDERGRADUATES THROUGH STATISTICAL CONSULTING," in *Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014)*, 2014, p. 4.

[27] T. Hooks and C. Malone, "Involving Undergraduates in Statistical Consulting," p. 11, 2012.

[28] B. Jersky, "STATISTICAL CONSULTING WITH UNDERGRADUATES – A COMMUNITY OUTREACH APPROACH," in *Proceedings of the Sixth International Conference on Teaching Statistics (ICOTS6, July, 2002)*, 2002, p. 5.

[29] K. Brundiers, A. Wiek, and C. L. Redman, "Real-world learning opportunities in sustainability: from classroom into the real world," *International Journal of Sustainability in Higher Education*, vol. 11, no. 4, pp. 308–324, Jan. 2010. [Online]. Available: https://doi.org/10.1108/14676371011077540

[30] S. Miller, "Collaborative Approaches Needed to Close the Big Data Skills Gap," Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 2458930, Apr. 2014. [Online]. Available: https://papers.ssrn.com/abstract=2458930

[31] Q. Cheng, F. Lopez, and A. Hadjixenofontos, "Integrating Introductory Data Science into Computer and Information Literacy through Collaborative Project-based Learning," in *2019 IEEE Frontiers in Education Conference (FIE)*, Oct. 2019, pp. 1–5.

[32] M. Micari and P. Pazos, "Beyond grades: improving college students' social-cognitive outcomes in STEM through a collaborative learning environment," *Learning Environments Research*, Jun. 2020. [Online]. Available: https://doi.org/10.1007/s10984-020-09325-y

[33] P. Ordóñez Franco, H. Ortiz Zuazaga, and et al., "Enhancing Undergraduate Education and Curriculum through an Interdisciplinary and Quantitative Initiative to Broaden Participation in Big Data," in *Proceedings of the 18th LACCEI International Multi-Conference for Engineering, Education, and Technology: Engineering, Integration, And Alliances for A Sustainable Development"*, 2020. [Online]. Available: http://laccei.org/LACCEI2020-VirtualEdition/meta/FP581.html

[34] J. Premo, A. Cavagnetto, and W. B. Davis, "Promoting collaborative classrooms: The impacts of interdependent cooperative learning on undergraduate interactions and achievement," *CBE—Life Sciences Education*, vol. 17, no. 2, p. ar32, 2018. [Online]. Available: https://doi.org/10.1187/cbe.17-08-0176

[35] E. J. Dallimore and T. J. Souza, "Consulting course design: Theoretical frameworks and pedagogical strategies," *Business Communication Quarterly*, vol. 65, no. 4, pp. 86–101, 2002. [Online]. Available: https://doi.org/10.1177/108056990206500408