Adopting Foundational Data Science Curriculum with Diverse Institutional Contexts

Vandana P. Janeja*
vjaneja@umbc.edu
University of Maryland, Baltimore County
MD, USA

Maria Sanchez
University of Maryland, Baltimore County
MD, USA
msanchez@umbc.edu

Yi Xuan Khoo
University of Maryland, Baltimore County
MD, France

Claudia Von Vacano
University of California, Berkeley
CA, USA

Lujie Karen Chen
University of Maryland, Baltimore County
MD, USA

ABSTRACT

The prevalence of data across all disciplines and the large workforce demand from industry has led to the rise in interest of data science courses. Educators are increasingly recognizing the value of building communities of practice, adapting and translating courses and programs that have been shown to be successful, and sharing lessons learned in increasing diversity in data science education. We describe and analyze our experiences translating a lower-division data science curriculum from one university, University of California, Berkeley, to another setting with very different student populations and institutional context, University of Maryland, Baltimore County (UMBC). We present our findings from student interviews across two semesters of the course offering at UMBC, specifically focusing on the challenges and positive experiences that the students had in the UMBC course. We highlight lessons learned to reflect on the existing large scale program at UC Berkeley, its adaptation and opportunities for increasing diversity in new settings. Our findings emphasize the importance of adapting courses and programs to existing curricula, student populations, cyberinfrastructure, and faculty and staff resources. Smaller class sizes open up the possibility of more individualized assignments, tailored to the majors, career interests, and social change motivations of diverse students. While students across institutional contexts may need varying degrees of support, we found that often students from diverse backgrounds, if engaged deeply, show significant enthusiasm for data science and its applications.

CCS CONCEPTS

• **Social and professional topics → Computational thinking**; **Model curricula**.

*Corresponding Author

SIGCSE 2024, March 20–23, 2024, Portland, OR, USA

ACM ISBN 979-8-4007-0423-9/24/03…$15.00
https://doi.org/10.1145/3626252.3630771

KEYWORDS

data science, computational science education, curriculum adoption, diversity in STEM education

ACM Reference Format:
Vandana P. Janeja, Maria Sanchez, Yi Xuan Khoo, Claudia Von Vacano, and Lujie Karen Chen. 2024. Adopting Foundational Data Science Curriculum with Diverse Institutional Contexts . In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE 2024), March 20–23, 2024, Portland, OR, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3626252.3630771

1 INTRODUCTION

According to the National Academy of Sciences and the Business Higher Education Forum, the information economy will require a larger proportion of the workforce to be data-enabled and data literate [12, 15]. Moreover, students are realizing the prevalence of data in their own lives and in their own disciplines earlier in their education experiences, leading to new career avenues [2]. This has prompted a surge of student interest in data science, overwhelming the current capacity of many institutions and forcing them to develop and teach new courses. Thus, the demand for proficiency in data science skills presents important challenges of scale for colleges and universities seeking to prepare their graduates for the modern workforce, especially in an environment of increasingly limited public investment in post-secondary education.

Making data science education available across institutions also requires broadening the participation of students who have been historically marginalized and not well represented through their lived experiences in STEM education, such as minoritized and marginalized communities, Black, Indigenous, and People of Color (BIPOC, including Latinx), women, and first-generation (first-gen) college students [13]. Moreover, because the data science work often involves dealing with data on humans, the field requires a diverse pool of practitioners who bring multiple perspectives to data analytics through lived experiences, and an understanding of the implications of their work on a society with rapidly changing demographics. Given that data science education is still relatively new on many college campuses, scientists and educators now have an unprecedented opportunity to achieve broader participation in data science through research-based strategies that can effectively serve all students.

Rather than creating new courses and educational programs from scratch to meet this demand, data science educators are developing a flourishing ecosystem of pedagogical and curricular materials that are shared and translated across institutions, from community colleges to 4-year colleges [1, 9, 10]. Yet, pedagogical practices and curricula cannot simply be copied from one setting to another. Instead, faculty and administrators wishing to adopt innovations from other institutions must make systematic and thoughtful efforts to translate what has worked in one setting and then adapt it into their own setting, fitting it into existing curricula and tailoring it to their own student populations' interests and needs. Also, replication of programs provides an opportunity to reflect on the program model and can reveal new insights regarding necessary changes to make existing programs more inclusive.

Many data science curricula have been proposed and implemented [3, 5, 7, 8, 11, 15, 16]. One early developer of a robust curriculum of data science education is University of California, Berkeley [3], which began its focus on scaling lower-division data science courses. The UC Berkeley model has already attracted national attention and interest in replication, with early indications suggesting that parts of the model can be successfully transferred to other institutions. A recent report from the National Academy of Sciences cites UC Berkeley's program as an exemplar in teaching and broadening participation in data science [15]. The UC Berkeley Data8 [1] course is built on the foundation of Jupyterhub that can be accessed from any browser, including from a phone, to interface with a python notebook based learning. This democratizes data science education with a low barrier to entry, making the course accessible to all types of institutions. The approach has sparked widespread interest as shown in the participation of faculty members from higher education institutions across the country in workshops on data science pedagogy using Data8 as a model, with over 500 participants each year since 2021[4].

In this paper, we analyze the process of adoption and translation of the UC Berkeley model for a lower-division data science foundational course to University of Maryland, Baltimore County (UMBC). Foundations of Data Science course [1] offered at UC Berkeley is an entry-level, four-unit course designed for any student with no prerequisites. It teaches critical concepts and skills in computer programming and statistical inference, in conjunction with hands-on analysis of real-world datasets, including economic data, document collections, geographical data, and social networks. UMBC adopted Data8 as an Information Systems Department [IS]] department course (IS 296: Foundations of Data Science) and has offered it six times since the fall of 2020. We adapted the course to our own institutional context with considerations for: (a) the student body and their backgrounds, (b) the number of credits/hours they can dedicate to studying per week with considerations for the additional workload that they might have to balance.

The two institutions had a different institutional context and diversity of the student body as shown in Table 1 based on 2023 data.

UMBC, with over 10,600 undergraduate students and with over 51% enrollment from Historically Underrepresented Groups (HUGs), is a Minority Serving Institute (MSI) that recently received the R1 Carnegie classification. UMBC is well known in the space of inclusive excellence and prides itself on the diversity of the campus emphasizing student success. UC Berkeley is California's flagship public research university with a student population of over 42,000 and over 350 different degree programs. Twenty-two percent of undergraduates are underrepresented minorities, 21 percent of undergraduates are transfer students, and 23 percent of first-year admits are first-generation students. We estimate that almost one-quarter of Berkeley's undergraduates take the Foundations course at some point in their college career. The aim of this study is to understand the student experience with the adaptation of the course Data8 from UC Berkeley to IS 296: Foundations of Data Science at UMBC. We draw parallels between the two institutions, while also foregrounding differences. Based on two semesters of student interviews, we discuss and evaluate qualitative interview data to identify lessons learned that have implications for other efforts at adaptation and translation of data science pedagogy and curricula. While we also conducted interviews with administrators and faculty, in this paper we mainly focus on the student experience. We particularly wanted to identify the challenges and positive experiences of the students in the class to inform possible future adaptations. We specifically address the research question: **What are the challenges encountered by the students in the foundations of data science class at UMBC? What are the positive experiences they had?**

Table 1: Institutional context: UCBerkeley and UMBC

Institution	UC Berkeley	UMBC
Carnegie classification	R1	R1
Public/Private	public	public
Total undergrad student body	30,853	10,625
Average Foundations class size	1500	25
Minority serving status	MSI (2023)	MSI
URM population	22%	51.8%
First-gen population	23%	25%
Female identifying undergrads	54%	47%

In the next sections, we outline the adoption, evaluate data collected through interviews and focus groups, and discuss lessons learned from these adaptations that reflect on the insights into what can be a successful approach to adopting the foundations of data science class.

2 CURRICULUM ADOPTION

Adoption: Since its adoption in the Fall of 2020, IS 296 has been offered six times. UMBC lowered the number of credits to three (as compared to four at UC Berkeley) with limited contact hours to fit the course into existing frameworks and include it on a pathway to existing degree requirements [17]. It was also important that the students did not have to pay for an additional course and increase time to graduation. As a result, UMBC adaptation made this course fulfill a programming requirement for one of the majors, a new business analytics certificate[17], and is also part of the X+CS effort [14].

Adapting course infrastructure: One key aspect of the course was the Python based Jupyter notebook hub. In general, setting up the JupyterHub infrastructure was non-trivial and needed ramp-up

time. The first run-through had challenges with versioning and interdependencies of Python libraries as well as security and authentication issues. However, once the infrastructure was set up and running, the students' experiences with authentication, accessing files and working in Jupyter notebooks, for the most part, were smooth, easy, and seamless. The trick to having a strong working infrastructure was getting the right expertise at the right time, and the right amount of time from experts, which was really based on a partnership with the UMBC office of information technology and lessons learned from the Foundations infrastructure at UC Berkeley. With the heavy reliance on Jupyter Notebooks for this Foundations course, it is imperative to have this established well in advance with sufficient testing in place.

Adapting the content structuring: We started with the same lab and project structures as the original Foundations course up to the midterm exam. A number of modifications were made to the content after the midterm. The focus on the curriculum shifted to a conceptual understanding of hypothesis testing, lighter on programming, while students honed programming skills learned in the first half of the semester (table and visualization functions) through project work. Some labs and homework were shortened or replaced with local contextual data and new in-class activities on data storytelling were introduced. The content from the last three chapters (regression, prediction, classification) were condensed and adapted to the local context. An ethics module, initially missing from Data8, was introduced and it included a guest lecture, a case study through the lens of ethical data life cycle and a follow up quiz. In-class labs were significantly modified allowing students to explore their own data, propose and answer their own curiosity driven questions. Students were given sample simulation code without having to write code from scratch. We also added new homeworks, projects, story-reading exercises where students are asked to provide a narrative for a given data visualization, and exploration of data science case studies to understand the type of research questions being addressed by the data. These changes helped increase flexibility for students to bring in their own lived experiences. One of the projects was a tool exploration where they could pick any data science tool (such as WEKA, Rapid Miner, Google Cloud based Machine learning etc.) and explore it with the datasets they had encountered in their case study. Another project was a custom group project where students picked their dataset, and delivered digital data stories in video format. Students were asked to submit weekly updates, each with several data insights. The cumulative set of data insights were assembled into data stories toward the end of the project. Each project was assigned a data science consultant (Data science scholars) and a storytelling consultant (Community College interns). Data Science scholars are near peer-mentors recruited from campus community with prior data science experience. The instructor and consultants provided feedback on the weekly update. The detailed adaptations are available [18].

3 STUDY METHODOLOGY

Data Collection: Semi-structured interviews and focus groups were the primary methods of data collection. The interviews consisted of students from the class and Data Science Scholars (DSS) who supported the class as peer mentors and teaching fellows that semester. The data science scholars also took this class in the prior semesters. These interviews were conducted in-person and via zoom and lasted approximately 60 minutes each. **Data Analysis:** Thematic analysis, a widely used qualitative analysis approach, was employed to analyze the data. We followed the stages of thematic analysis by Braun and Clarke [6] to analyze the data. The interviews and focus group sessions were video-recorded and transcribed verbatim using a professional transcription service, Rev.com. The transcripts were then imported into MaxQDA TeamCloud for analysis. Two researchers, trained in qualitative research methods, coded the data. Regular meetings across the team were held to discuss coding decisions and address any discrepancies that arose. Initially, the coding process focused on identifying key themes that emerged from the data. Once the initial coding was completed, the researchers reviewed the codes and identified overarching themes that reflected the experiences of the participants. The researchers engaged in a collaborative process to develop a narrative that captured the key themes and experiences of the participants. We then grouped the thematic codes into what we call categorized code trees (Figures 1 and 2). Overall, this cyclical process of coding, reviewing, and analysis in MaxQDA TeamCloud allowed for a thorough exploration of the experiences of students and provided insights for improving support services to better meet the needs of all students including those from HUGs.

4 STUDY DATA AND RESULTS

We assess the adaptation through the lens of the research question in terms of the classroom experience in IS 296. We outline the demographics of the student interview participants and then evaluate these interviews through the codes created to assess the question. We then identify emerging themes and highlight quotes exemplifying these themes. We note that there were additional interviews with faculty and administrators; in this paper we mainly focus on the student experience.

4.1 Demographic distribution

The data we evaluated consists of interviews of 16 participants (13 students and 3 Data Science Scholars who also took the class) across two semesters. Table 2 shows the demographic categories of the students that participated in the study (some individuals did not provide this information).

4.2 Results

4.2.1 Summary of Key themes emerging. Table 3 summarizes the overall themes that emerged across the interview analysis. These Key themes reflect the positive experiences and challenges along with reflections across identity, background and experiences of inclusion and exclusion. In the following section we discuss these Key themes with the lens of categorized code trees.

We saw that peer interaction and instructor support was important in our adaptation of the course. We also saw that students did not feel a sense of exclusion despite the lack of programming experiences. It was indeed clear that students did have programming difficulties and wanted a slower pace of the class. This is interesting to note since there was some dichotomy across backgrounds where some students wanted more resources and challenges in the class due to better preparation. This exemplifies the diverse student body

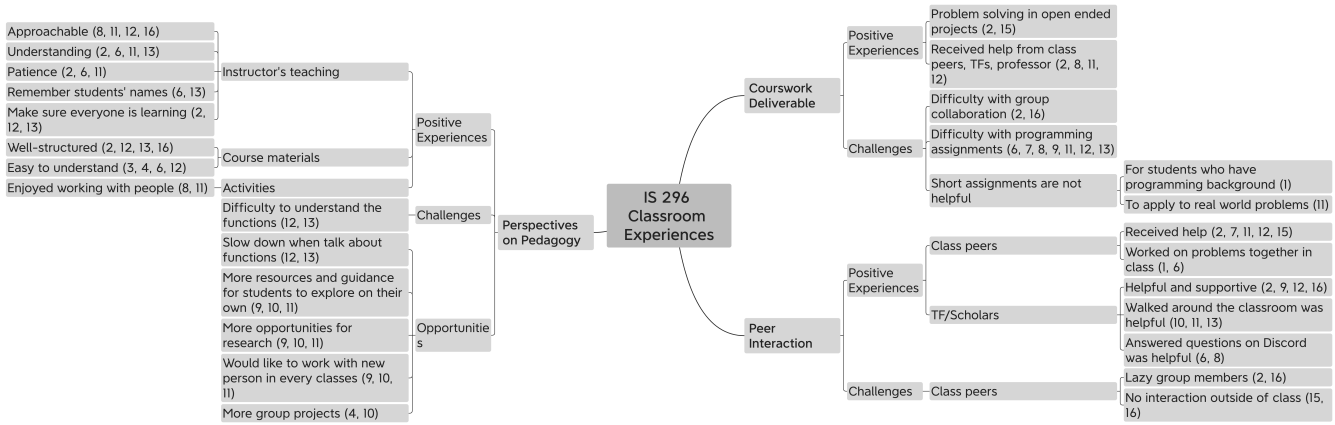

Figure 1: IS 296 Classroom experiences

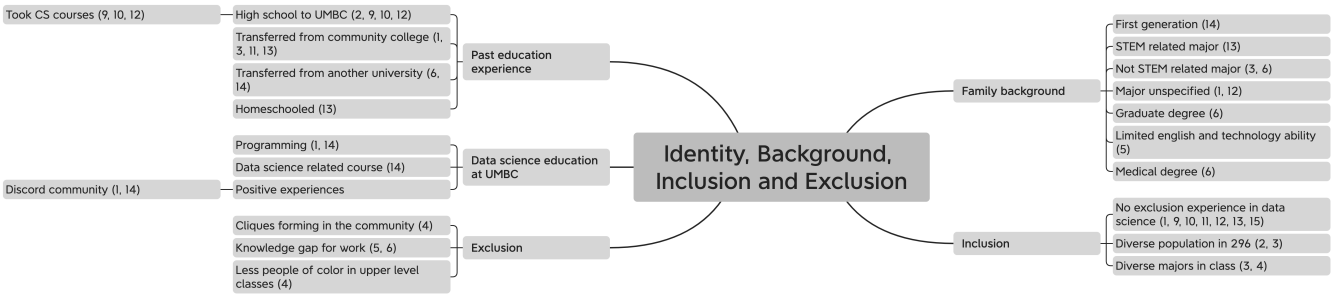

Figure 2: General Findings

Table 2: Demographic distribution of students interviewed

Ethnicity	Gender	First Gen	Major	Transfer	Year
Black: 5 White: 3 Asian: 5 Mixed: 3	Male: 9 Female: 7	Yes: 2 No: 6 No response: 8	Computing: 5 Non-Computing: 10 Undecided: 1	Yes: 7 No: 8 No response: 1	Freshman: 4 Sophomore: 1 Juniors: 6 Seniors: 2 No Response: 3

and the challenges to address balancing the diverse experiences in the classroom.

4.2.2 RQ1: IS 296 curriculum adoption. Figure 1 shows the categorized code tree created to analyze the data and assess the question: What are the challenges encountered by the students in the foundations class? What are the positive experiences they had?

The overall key themes are shown in Table 3. Below we provide quotes from the interviews to illustrate the analysis supporting the overall themes.

Support in class: We see positive experiences where the students appreciated help from the instructor, peers and teaching fellows as the coursework was delivered. Additionally, students felt supported by their own peers in the class when they worked on assignments. For example participant 11 noted about teaching fellows *"...I was just doing my thing and then he passed by and he asked me if I had a question. So that was very helpful. So I actually had a question because I did not know how to do something and he walked me through it..."* Similarly participant 6 noted for the instructor that *"...I think [they] really makes the class itself very easy to learn."* Participant 11 also noted about the instructor *"If you don't get it, [they are] willing to stay and help you until you get it."* The instructor's availability and personalized approach was very helpful in maintaining the interest in the class. It is clear that the smaller class and the presence of peer interaction has a positive impact on the students. Similarly the support provided by the instructor has a major impact on the students. This is an important consideration as the adaptation is scaled in a diverse setting. Students also noted that they enjoyed working in a class with different types of people and embraced the idea of working with peers, for example participant 11 said *"So it helped us exchange contact with somebody so you know that, hey, if*

Table 3: Key Themes from code trees

Observations from the Research Code Trees	Key Themes
RQ 1 A: What are the positive experiences encountered by the students in the foundations class?	Peer interaction, diversity of disciplines, instructor approachability, flexibility, open ended group projects
RQ 1 B: What are the challenges encountered by the students in the foundations class?	Some dichotomy on difficulty with programming assignments, pace of the class, need for more resources for students who may be advanced
Reflections on Identity Background. Inclusion and Exclusion	Sense of inclusion in the class, despite diversity in background and preparation the overall experiences are positive, support networks and peer interaction play an important role in student experiences. Prior computing experience makes a difference in the programming experience in the classroom. Majors outside traditional computing (such as Business technology majors), freshman and juniors had difficulties with the programming experience in the class.

I'm not sure how to do this homework, I can take that person to help me out and stuff. So it's pretty nice to work with people around"

Flexibility and Structuring: Students felt that working on open ended projects helped them explore, as expressed by participant 15, *"...I like about our class the most is how open ended especially, we have a big project that is due.....but just having the openness to pick the data sets that we want to analyze, and work in because there's data sets for basically every different subject".* Students felt that the course materials were well structured and easy to understand. This supports the idea of adoption since UCB has evolved the material a lot over several iterations. The instructor has started varying some datasets in more advanced notebooks but at the same time the existing jupyter notebooks for the foundational pieces seemed to help the students' understanding. For example, participant 12 noted *"I mean, at the beginning of the week, [they] basically gives us the plan that we're going to do this whole week, and then, she goes over it, and then, we followed through....[they are] pretty structured when it comes to education"*

We next discuss challenges and opportunities that the IS 296 students faced as shown in Table 4 such as difficulty with programming assignments, pace of the class, need for balancing across student backgrounds.

Curriculum and deliverables: Several students reported difficulty with programming challenges. However students with an existing background in programming wanted more challenging real world problems and did not find short assignments helpful. This reflects on the levels of students in the class which is a particular challenge in a diverse setting where students come from many types of experiences and backgrounds. For example, depending on where the students come from, there may be students who never had programming in high school and others who may have had programming experiences from high school. Participant 8 noted *"I was completely unfamiliar with coding and creating something of my own, so I was interested to learn how everything that I use daily works".* Students recommended slowing down on some complex topics (for example, functions). This also reflects on the depth vs breadth of topics in the original Data8 curriculum, which may not translate well to all types of institutions. Thus, there is a need to balance the number of topics covered in lieu of slowing down and spending more time on others making sure students gain confidence as they learn, as expressed by participant 13, *"...Actually, maybe just for me, I think she could slow down a little when it comes to talking about a function, or whatever it is that we're talking about."* and participant 12, *"....I wouldn't say it's fast paced, but yeah, it's pretty fast paced..".* Students were interested in more resources to explore on their own and also opportunities for research. For example participant 10 noted *"But yeah, instead of just having a worksheet, filling it out, submitting it, maybe a little more research, kind of using what you learned as well and then sharing the class."*

4.2.3 *Reflections on Identity Background, Inclusion and Exclusion.*

Figure 2 shows the categorized code tree for Identity Background, Inclusion and Exclusion. We found key themes emerging as shown in Table 4. We see that while there is no clear trend in programming experiences there are some common threads emerging. Transfer students had difficulty with programming. Background in computing seems essential (Business oriented technology majors had difficulty, undecided also had difficulty). The class design needs to have a serious consideration for non tech majors and no prerequisites. For example participants 7 and 8 both from these majors noted a lack of prior preparation. For some students, having difficulties with programming did not seem to affect their sense of belonging in the class (did not experience exclusion). Similarly these students who had difficulties also had an overlap with students that found the instructor and peer interaction helpful. For example participant 13 noted *" since this is my first computer, anything class, I was kind of weary....[they are] very inclusive. I feel like that's how I knew your name, it's because she's always saying it or somebody, [they are] really good at that, and making everybody feel like, either one, they have to participate or they should."* Creating an inclusive environment and educational background both play a clear role in their overall experience in the class. Students recognize good experiences even if they struggle with the academic experience in the class. One of the African American students, participant 4 (out of five in the population), expressed a sense of exclusion in the wider STEM community and lack of students of color in upper elective classes noting *"Well, I'd say that the harder my STEM classes are being, the less people that look like me, I've seen. So I guess people, when they look at you, they don't assume that you can do those kind of things or you can think that way. So you have to go and look for those opportunities because they won't come to you, and I think that's the thing I learned*

pretty early on". On the other hand several female students had positive peer interactions and also highlighted the support network such as Center for Women In Technology (CWIT). We partnered with CWIT to provide professional development and mentoring support for the peer mentors as well. For example participant 3 noted *"a networking event I went to for CWIT, there were tables of just minorities gathering together and what not. And it just happens at CWIT, [they have] a lot of compassion for minorities, and support for them, and our struggles".*

4.2.4 Summary of reflections on the Adaptation of IS 296. Overall we believe that the adaptation was successful in aligning to the UMBC context. We noticed that over the semesters of the course offering the overall grade distribution has improved. The rate of DFW grades (where students received a D, Failed or Withdrew) was 19.2% (F22) and 11.5% (Sp23), and average grade was 3.08 (F22) and 3.33 (Sp23). In general we saw an improvement across semesters in the DFW rate and average grades. Our interview process clearly exemplified the benefits of working in a smaller class as compared to the Data8 class which is a much larger setting. Several aspects of the curriculum such as the infrastructure and the Jupyter Notebooks are easily adaptable where others needed much more contextualization.

Here we share lessons learned through adopting the Foundations course, some of which may inform discussions around diversity and nurturing a diverse data science workforce through a balance of structure and flexibility, large and small scales.

Key Modifications: Here, we highlight two major areas which may require investigation from others looking to adopt the Foundations course - (a) we shortened the labs and lectures to fit our institutions' credit model and style of teaching and to better align with students' backgrounds; (b) smaller class sizes allowed for experimentation in changed content and labs, making it possible to have more flexible assignments and projects that could be graded in an individualized manner. These assignments helped take advantage of the different backgrounds, different interest areas, and diverse mindsets of the students by encouraging and supporting diversity in thought and dialogue. The detailed adaptations are available [18].

Challenges: We discuss two aspects of the adaptation that present challenges. (a) Content specific challenges: The first set of challenges involves content and delivery. Some topics were not well explained or supported by datasets, such as probability, which was laid out in a silo without relevance to future topics. In some topics we felt that there was an assumption of the student's preparation. This can get aggravated with a mix of types of students in the class, including students coming from different types of high schools (with or without computing curriculum), freshman, and transfers with different levels of preparation. We also found some topics needed to be better explained, such as histograms and randomization, while more computation intensive topics, such as histograms, p-values, confidence intervals, needed more practice problems. (b) Pace of the class: The second set of challenges involved the technical nature of the course, where some students found the computing and statistical concepts difficult and the pace of the course fast. This is a reflection of the institutional context as well. A possible solution is to offer a short module where students could get exposure to some computing and statistical topics before taking the foundations course. It is noteworthy that both UCB and UMBC have reached this observation even with different institutional contexts and levels of student preparation. Another possibility is to offer an alternative non-computing intensive course that focuses more on social science and ethics issues related to data science. This will allow for variations of the course for different types of student populations while keeping the essence of the foundations of data science intact.

Reflections for UC Berkeley: The experiences and adaptations made by UMBC carry lessons for UC Berkeley's data science program as well, particularly with regard to best serving students from groups traditionally underrepresented in STEM fields, specifically: (a), course material on topics related to human contexts and ethics in data science can serve as an entry point for a wider range of students by demonstrating the relevance of technical material to contemporary social problems and public policies; (b) opportunities for more individualized assignments and projects can also motivate students to continue the pursuit of their own interests within data science; (c) need for explicit and ongoing discussions of the importance of diversity in data science may help to provide students from marginalized groups with a greater sense of belonging and identification with data science as a field: and (d) the Data Science Scholars program along with the near-peer mentors can be geared towards supporting a more individualized approach.

5 CONCLUSIONS

In this experience report, we have described and analyzed our efforts to adapt and translate a lower-division data science course from a large public R1 university (UC Berkeley) to an R1 university serving a diverse population (UMBC). Our findings highlight the importance of beginning small before attempting to scale translated programs and the need to adapt courses and programs to existing curricula, student populations, cyberinfrastructure, and faculty and staff resources. Smaller class sizes open up the possibility of more individualized assignments that are tailored to the majors, career interests, and social change motivations of students who may bring in lived experiences, and while some students may need support to ramp up, they also show significant enthusiasm for data science and its applications. In this paper we analyzed student interviews across two semesters. We hope to extend this work to include faculty and administrator perspectives. We also plan to analyze findings from survey data across six semesters of course offerings at UMBC. Additionally, we plan to include other research questions about data science career pathways. Lastly, we plan to study the adaptation in comparison across multiple institutions.

ACKNOWLEDGMENTS

This work is supported by NSF grant # 1915714

REFERENCES

[1] UC Berkeley. 2015. Foundations of Data Science. http://data8.org/.
[2] UC Berkeley. 2018. Translational Data Science workshop reflections. https://data.berkeley.edu/news/translational-data-science-workshop-reflections-featured-nature.
[3] UC Berkeley. 2022. Berkeley Computing, Data Science, and Society Curriculum Overview. https://data.berkeley.edu/curriculum-overview.
[4] UC Berkeley. 2023. National Workshop on Data Science Education. https://data.berkeley.edu/2023workshop
[5] Ismail Bile Hassan, Thanaa Ghanem, David Jacobson, Simon Jin, Katherine Johnson, Dalia Sulieman, and Wei Wei. 2021. Data Science Curriculum Design: A Case

Study. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (Virtual Event, USA) *(SIGCSE '21)*. Association for Computing Machinery, New York, NY, USA, 529–534. https://doi.org/10.1145/3408877.3432443

[6] Victoria Clarke, Virginia Braun, and Nikki Hayfield. 2015. Thematic analysis. *Qualitative psychology: A practical guide to research methods* 3 (2015), 222–248.

[7] Richard D De Veaux, Mahesh Agarwal, Maia Averett, Benjamin S Baumer, Andrew Bray, Thomas C Bressoud, Lance Bryant, Lei Z Cheng, Amanda Francis, Robert Gould, et al. 2017. Curriculum guidelines for undergraduate programs in data science. *Annu Rev Stat Appl* 4 (2017), 15–30.

[8] Yuri Demchenko, Adam Belloum, Wouter Los, Tomasz Wiktorski, Andrea Manieri, Holger Brocks, Jana Becker, Dominic Heutelbeck, Matthias Hemmje, and Steve Brewer. 2016. EDISON Data Science Framework: A Foundation for Building Data Science Profession for Research and Industry. In *2016 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*. 620–626. https://doi.org/10.1109/CloudCom.2016.0107

[9] Thomas Donoghue, Bradley Voytek, and Shannon Ellis. 2022. Course Materials for Data Science in Practice. *Journal of Open Source Education* 5, 51 (2022), 121. https://doi.org/10.21105/jose.00121

[10] Thomas Donoghue, Bradley Voytek, and Shannon E. Ellis. 2021. Teaching Creative and Practical Data Science at Scale. *Journal of Statistics and Data Science Education* 29, sup1 (2021), S27–S39. https://doi.org/10.1080/10691898.2020.1860725 arXiv:https://doi.org/10.1080/10691898.2020.1860725

[11] ACM Data Science Task Force. 2021. *Computing Competencies for Undergraduate Data Science Curricula.* Association for Computing Machinery, New York, NY, USA.

[12] Business-Higher Education Forum. 2017. *Investing in America's Data Science and Analytics Talent: The Case for Action.* Technical Report. PwC and the Business-Higher Education Forum.

[13] National Science Foundation. 2013. *Committee on Equal Opportunities in Science and Engineering (CEOSE) Biennial Report to Congress, Broadening Participation in America's STEM Workforce.* National Science Foundation, Arlington, VA.

[14] Susan Mitchell, Katharine Cole, and Anupam Joshi. 2020. X+ CS: A Computing Pathway for Non-Computer Science Majors. In *ASEE Mid Atlantic Section Spring Conference, 2020.*

[15] National Academies of Sciences Engineering, Medicine, et al. 2018. Data Science for Undergraduates: Opportunities and Options. (2018). https://doi.org/10.17226/25104

[16] Aimee Schwab-McCoy, Catherine M. Baker, and Rebecca E. Gasper. 2021. Data Science in 2020: Computing, Curricula, and Challenges for the Next 10 Years. *Journal of Statistics and Data Science Education* 29, sup1 (2021), S40–S50. https://doi.org/10.1080/10691898.2020.1851159 arXiv:https://doi.org/10.1080/10691898.2020.1851159

[17] Baltimore County University of Maryland. [n. d.]. Bachelor of Arts in Business Technology Administration. https://informationsystems.umbc.edu/home/undergraduate-programs/undergraduate-degree-programs/bachelor-of-arts-in-business-technology-administration

[18] Baltimore County University of Maryland. 2018. IS 296 Adaptations. https://tinyurl.com/296adapt.