The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
ASEE
Paper ID #43073

An Online Interdisciplinary Professional Master's Program in Translational Data Analytics

Dr. Emily Nutwell, The Ohio State University

Dr. Emily Nutwell is currently serving as the Program Director of the Masters in Translational Data Analytics at the Ohio State University. This applied program, designed for working professionals, focuses on the foundation of data analysis, computing, machine learning, data visualization, and information design. Prior to joining Ohio State, Dr. Nutwell worked at Honda R&D Americas for close to twenty years as a vehicle crash analysts specializing in computational techniques. She holds a BS in mechanical engineering, MA in educational studies, and a PhD in Engineering Education where her research focuses on digital learning environments for the STEM workforce.

Thomas Bihari, The Ohio State University
Thomas Metzger, The Ohio State University

An Online Interdisciplinary Professional Master's Program in Translational Data Analytics

Abstract

This paper describes an interdisciplinary data analytics professional master's program which includes courses from the disciplines of computer science, statistics, and design. The online curriculum structure specifically addresses the needs of working professionals with little to no prior data science, computing, or math background. Courses use both synchronous and asynchronous delivery methods to maximize learner flexibility while providing opportunities to engage in real time with instructors and peers. All courses emphasize projects to provide opportunities for learners to apply courses concepts to real-world problems. A terminal 2-semester capstone course incorporates all three disciplines into a final culminating team project. This paper will focus on the conceptualization of the computer science (CS) portion of the curriculum. As an applied master's program, much of the CS curriculum takes inspiration from industry frameworks such as CRISP-DM and Agile project management to contextualize concepts. The curriculum incorporates design and design thinking concepts to emphasize creative problem-solving skills and the importance of data storytelling.

There is a need for educators to understand how to develop a curriculum for working professionals which introduces novice programmers to 1) core data and computational concepts; 2) tools and techniques; 3) data-driven problem-solving workflows; and 4) data storytelling. This paper presents these four "swim lanes" to define a framework for describing a cohesive interdisciplinary curricular experience for an applied master's program.

Through reflection, the authors conclude that learners initially struggle with new concepts, but with sufficient support, they successfully learn and apply data science and computer science concepts in both didactic and experiential settings. Students appreciate the need to successfully communicate with data and be effective data storytellers but will often feel frustrated that data storytelling skills are not "real data science." An analysis of LinkedIn profiles indicates that over 60% of graduated learners secured new employment in data careers since starting the program. To build on this success, further curriculum development should more explicitly connect fundamental data science concepts and broader concepts such as creative problem-solving and data storytelling.

Keywords

Graduate education, data analytics, distance learning, life-long learning, adult learning

1. Introduction

We are living in an era where the Volume, Velocity, Veracity, Variety, and Value of data is being harnessed in unprecedented ways to support decision making in research and business [1], [2], [3]. Working professionals are recognizing compelling opportunities for career growth with the job outlook in data science and analytics far above the average of 5% with a projected growth rate of 36% for 2021-2031 [4]. Both graduate and undergraduate programming are needed to address this growth, and graduate programs designed for nontraditional students provide opportunities for them to earn valuable formal credentials in data science and analytics [5], [6], [7].

In recent years, universities have been establishing data science masters' programs to meet this growing demand in the workplace, and many of them are well suited for working individuals by accommodating online learning modalities and part-time enrollment [8], [9], [10]. Since data touches every profession, a challenge is to develop effective computer science graduate programming for individuals who do not have traditional STEM experience, since most individuals graduate from non-STEM programs [3], [11].

Work on establishing this interdisciplinary graduate program began late 2016, and the program launched in Autumn 2020. As part of establishing a graduate program designed for working professionals in data analytics, local and regional partners were gathered for a series of meetings so that employers from industry, government, and non-profit organizations could provide direct feedback on the structure of program curriculum. The results of these meetings created a draft curriculum which included three broad areas of study: computer science and statistics to establish data science foundational concepts, and design to emphasize the importance of data storytelling [12], [13].

The inclusion of a design course sequence is incorporated into the program to emphasize data visualization and data storytelling. This includes providing learners with a general understanding of design principles that contribute to and enhance readability and aesthetics to effectively communicate insights from the data analysis. User experience and human computer interaction is included to discuss user understanding and stakeholder audiences for data visualizations.

This paper will focus on the computer science track in the interdisciplinary curriculum, where the goal is to provide a foundational presentation of computer science principles within the context of an interdisciplinary graduate program. The courses are designed to support learners in identifying common data structures and sources, using information technology and relevant programming environments to convey and retrieve information, and identifying processes and mechanisms commonly used to retrieve, assess, re-engineer, manipulate, and visualize data. The diverse backgrounds of the learners make this an interesting challenge for curriculum designers. How can a professional master's degree successfully introduce foundational computer science concepts for adult learners from diverse backgrounds?

2. Professional Learners

To date, the program has graduated two cohorts of adult learners, with 19 individuals in the first cohort, and 14 in the second cohort. The learners enrolled in the program were mostly novice programmers, with either limited or no prior programming experience. Although some learners did have informal training in basic programming skills, the learners mostly earned bachelor's degrees whose curriculum did not have a strong emphasis on computer science skills. Those that were STEM focused were primarily in allied health fields and social sciences. A breakdown of earned undergraduate degrees of learners who have successfully graduated from the program is shown in Fig. 1.

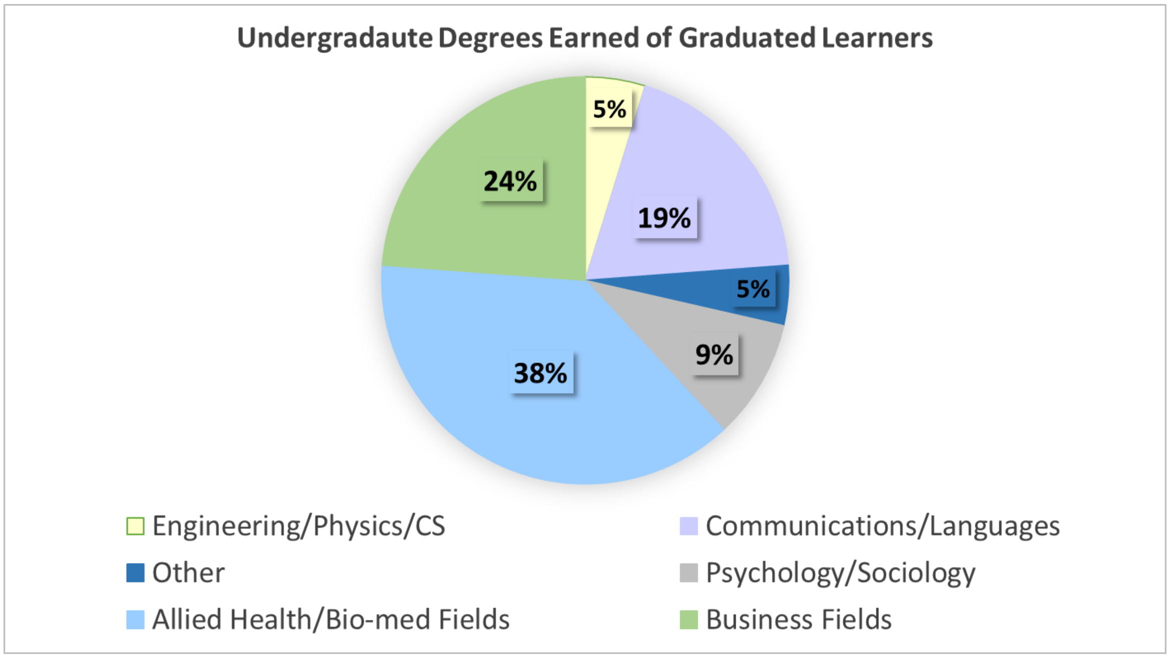

Fig. 1: Breakdown of earned undergraduate degrees of graduated learners.

Given the learner audience, the curriculum of the computer science curriculum needed to be approached considering that these learners are new to computer science concepts. Although this is a graduate program, the purpose of this program is to provide rigorous coursework to onboard learners to data science and analytics methods. Rather than a program that provides a deeper dive into a specific topic area, the goal of this professional master's degree is to provide an opportunity for individuals with expertise in diverse areas such as business, education, and allied health fields to incorporate data analytics into their professional practice.

3. Curriculum Overview

As a professional master's degree, the curriculum focuses on practical skill development to support adults considering career advancement through development of data analytics skills. The interdisciplinary curriculum is structured to integrate across data analytics disciplines

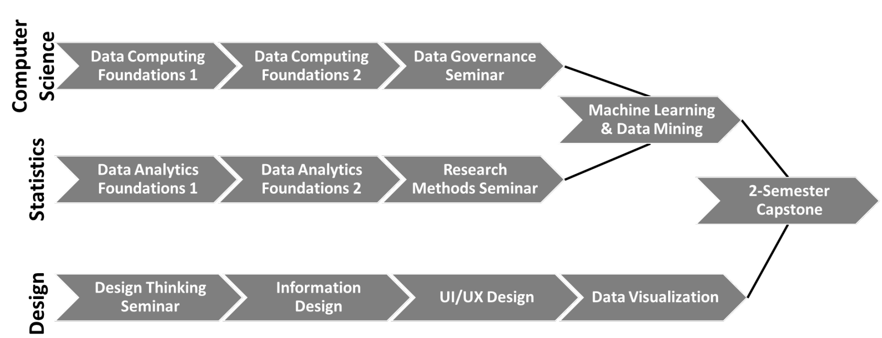

Fig. 2: Schematic of interdisciplinary data science professional master's degree

while being scoped to function effectively within the full program curriculum [14], [15]. A schematic of the curriculum is shown in Fig. 2.

The individual courses are structured in weekly modules that cover specific topics. Generally, each weekly module consists of:

- An assigned reading (e.g., from an online textbook or provided directly in the Learning Management System - all materials are provided to the learners free of charge).
- A recorded lecture and slide deck.
- Knowledge check quizzes (ungraded) and/or discussion posts
- A graded homework assignment.

The courses are online-asynchronous, and there is an optional weekly synchronous session at which learners can work with the instructor and with each other. These sessions are recorded and posted to the course for learners who cannot attend synchronously. These synchronous sessions are popular with the learners, and despite being optional, learners typically make the effort to attend weekly.

The courses are hosted on the university's standard Learning Management System (LMS). Microsoft Teams is used in some courses as a collaboration tool to support the online classroom. In addition to the synchronous sessions and chats, the instructors and teaching assistants hold weekly office hours or office hours by appointment to provide additional support to the learners.

Table 1: Overview of CS Curriculum for a Translation Data Analytics Master's Program

<table>
<tr><th>Course Name</th><th>Credits</th><th>Description</th></tr>
<tr><td>Big Data Computing Foundations 1</td><td>3</td><td>Locate, scrape, process and clean data to develop practical workflows which extract useful information and create usable representations. Focuses on the use of Python.</td></tr>
<tr><td>Big Data Computing Foundations 2</td><td>3</td><td>Create scalable data organizations and access to high-performance computing workflows for big data.</td></tr>
<tr><td>Data Governance</td><td>1</td><td>Presents practical elements of good data governance, privacy and data security through the use of case studies.</td></tr>
<tr><td>Learning & Mining</td><td>3</td><td>Split into two modules focusing on (1) practical and scalable data mining and (2) scalable machine learning.</td></tr>
<tr><td>Capstone 1</td><td>3</td><td rowspan="2">Culminating experience through direct engagement with community partners who will formulate challenge questions and provide data. Emphasis on teamwork, translational competency, interpretation of domain specific results, and data storytelling.</td></tr>
<tr><td>Capstone 2</td><td>3</td></tr>
</table>

Learners progress through the courses week by week. We encourage the learners to align to the weekly cadence for several reasons:

- By having all learners moving through the curriculum at the same rate, and cover the same topics together, common questions can be identified and addressed.
- Discussions in MS Teams are productive.
- Learners help each other.

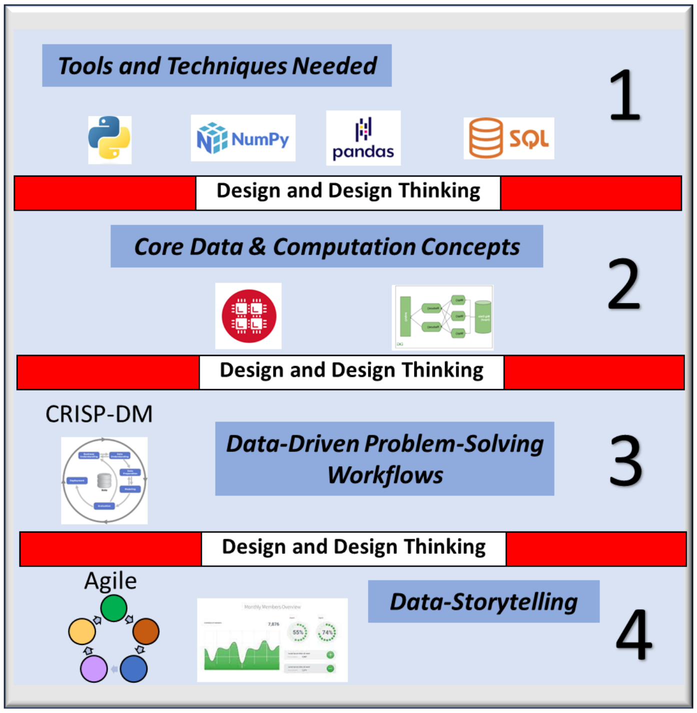

Figure 3: Concept for Computer Science curriculum as part of an interdisciplinary master's program.

The computer science courses emphasize case studies and projects to incorporate problem-solving and application of design concepts such as information design and data visualization. Fig. 3 depicts the four "swim lanes" of the curricular concept.

In addition to specific courses covering design topics such as information design, UI/UX concepts, data visualization, and design thinking, the curriculum emphasizes design and design thinking throughout the curriculum with an emphasis on creative problem solving and data visualization, by integrating these concepts throughout the computer science and statistics curriculum.

3.1 Foundational Computing Courses

The Big Data Computing Foundations course sequence spans two semesters. The Foundations 1 course is generally taken in the first semester of the program. The expectation is that the incoming learners have no formal computer science skills. In practice, some learners may have very strong skills in programming languages, database management and related concepts through using those skills in their daily jobs. Other learners may be new to some or all these concepts. This diversity of skills requires personal, tailored support. In addition to the standard coursework and direct support from the instructor and teaching assistant, the program offers supplemental programing such as tutoring to provide the needed support.

During the first portion of the semester, learners work on individual assignments. In the second portion, some assignments are executed by 2-3 person teams. The first course concentrates on bringing the learners to a consistent skill level in several areas, including:

- Programming in Python, using Jupyter Notebooks.
- Locating and accessing data sources.
- Manipulating data, including the use of NumPy, Pandas, and related tools.

- Building effective data analysis workflows.
- Authoring useful summary reports of the analyses.

While Python is used in the Computer Science courses, R is introduced in Data Analytics Foundations 1 and 2, both taught by faculty from the Department of Statistics. These programming tools are incorporated in conjunction with Python as the main coding tools for the course Practical Learning and Mining for Big Data. This gives the learners the opportunity to become competent in both languages and to see the pros and cons of each. This approach creates synergies across the overall interdisciplinary curriculum.

The Google Colab web-based environment is used for the programming exercises. This minimizes the need for the learners to install and maintain Python / Jupyter environments on their personal computers and provides an easy avenue to creation of the learners' first "Hello World" program.

The second foundation course builds on the concepts learned in Foundations 1. The expectation is that all learners at this point have a solid understanding in basic data-centric programming. In practice, some learners who were new to the concepts covered in Foundations 1 may need additional support, which the program continues to provide.

The Tools and Techniques thread, Lane 1 in Fig.3, in this course build on that of Foundations 1. We introduce SQL, and dive more deeply into use of Pandas. We pivot to using the Ohio Supercomputer Center's (OSC) supercomputers[16], instead of Colab. The programming interface is similar to Colab (web-based JupyterLab Python environment), and learners typically make the pivot without significant issues. Using OSC allows learners to process larger datasets, and to do some programming in Spark (Pyspark). The learners have the opportunity to process data with a variety of tools (SQL, Pandas, Spark SQL, Spark Functional libraries) and consider the pros and cons of particular tools.

The Core Data and Computation Concepts thread in this course addresses:

- Data Characteristics - (e.g., nominal, ordinal…) and what operations can be done on different types of data.
- Data Models - (e.g., relational, graph, column…) and their pros and cons for use in solving particular problems.
- Computation and Complexity – Why are some problems so hard to solve (big O notation)?
- Parallel Processing and Cloud Computing – How can we apply parallelism to deal with computational complexity?
- Frameworks: Hadoop, Map/Reduce, Spark – How can we apply frameworks to simplify application of parallelism?

The main goal of the Data-Driven Problem-Solving Workflows thread in this course is that learners understand the value and the characteristics of defined workflows, and that these workflows help us solve real-world problems. As discussed above, we use CRISP-DM (Cross-Industry Standard Process for Data Mining) [17] as a framework throughout the course. CRISP-DM is not the only such process – in fact, there are many [18], and we discuss this with the

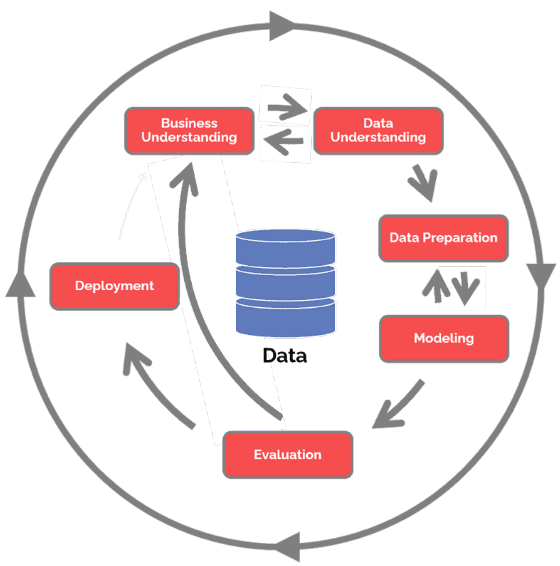

Figure 4: Depiction of the CRISP-DM process model. From https://www.datascience-pm.com/crisp-dm-2/

learners. However, CRISP-DM is widely known and used, and there is a wealth of free, available resources to support the learners in understanding the value of such frameworks and how to use them. Fig. 4 shows a common diagram describing the overall CRISP-DM workflow [19].

Throughout the term, the learners work incrementally through the six CRISP-DM steps. We structure a subset of the assignments throughout the semester to follow the CRISP-DM steps. For example, an early homework assignment requires the learners to define a "Business Understanding" step for a hypothetical real-world scenario. Later in the semester, a significant real-world problem and dataset is provided, and learners build up a Jupyter Notebook incrementally over multiple assignments, resulting in a large, end-to-end work product suitable for incorporation in a learner's portfolio. This thread generally gets positive feedback from the learners. Many of the learners are involved, or wish to be involved, in data-related problem-solving in their daily jobs. We encourage the learners to share stories about the kinds of situations they encounter in their jobs and to share their thoughts on how to improve the outcomes they observe. Occasionally, learners will share their frustration with the "non-programming" steps in CRISP-DM, for example asking, "When do we get to start writing more Python code?" when working through the Business Understanding step. However, typically the other activities that are running in parallel with this thread provide some of the desired programming practice.

The Data Governance seminar is a one credit hour course that provides learners with a deeper understanding of the meaning and scope of fundamental data governance principles, the issues that arise when issues arise around data-related cost and risk, and an overview of tools and techniques needed to provide sound data governance for an organization.

3.2 Data Learning and Mining Course

The Practical Learning and Mining for Big Data is a three-credit hour course co-taught between the Department of Statistics and the Department of Computer Science and Engineering. Content primarily addresses Lane 2 of Fig. 3, including considerations in data exploration and machine learning from a high-level perspective. The course aims to balance mathematical detail with a level of rigor appropriate for learners of our background. Unsupervised and supervised learning methods, including distance-based clustering algorithms, predictive regression modeling, neural networks, and tree-based methods, are emphasized. Each topic is motivated by the analysis of a canonical, open-source data set such as those found within the UC Irvine Machine Learning

Repository [20]. Data Storytelling is a method to communicate relevant information to stakeholders in an effective manner using data visualizations along with written and/or oral communication. The challenge for the analyst is to decide which information to include, and what to discard to focus attention appropriately to convey meaning with data [13], [21].
This course emphasizes practicality and interpretations of each method presented and assesses learner performance and understanding by evaluating learner-generated technical reports. In the course's final project, learners are encouraged to select a data set relevant to their personal area of work or interest. Past examples include projects relating to surgical outcomes, campus-wide energy efficiency, financial and banking outcomes, and agricultural yield. In keeping with Lane 4 goals, learners formulate a valid research question and use at least two of the methods discussed throughout the course to analyze their question.

3.3 2-Semester Capstone Course Sequence

The capstone is a 2-semester course sequence where learners work with project sponsors on non-trivial problems. The first few weeks of the first semester capstone course addresses Lane 3 of Fig. 3 by building off the introduction of the CRISP-DM workflow from the second foundational course. The first half of this course takes a deeper dive into the first two stages, namely the development of the business understanding and data understanding of the project. The learners are working with project sponsors from various community, government, or industry partners and need time to understand the contextual nuances of their projects. The data sets are complex, and problems are ill-defined. The CRISP-DM process model provides a framework to define project goals that are well aligned with the sponsors' true objectives [22], [23], [24]

Since learners are working in teams, the course also explicitly discusses strategies for effectively working in teams and defining inclusive teams. The learners are working professionals, and it is safe to assume that most of them have worked in teams in the past. However, presenting principles and strategies for effective teamwork clarifies expectations, prepares them to work collaboratively, and gives them the tools to develop their team effectively [25], [26]. This explicit introduction to teamwork principles addresses Lane 4, Data Storytelling, since effective teamwork, collaboration, and communication are all required elements of data storytelling.

The goal of the first semester is to establish project goals (business understanding), generate a first draft of a data dictionary which is aligned the project goals (data understanding), and conduct an exploratory data analysis (data understanding). Based on this work, the learners then create a project plan for their capstone project.

The first semester also introduces an Agile project management flow adapted for data science. Rather than a traditional waterfall approach which requires significant upfront planning, Agile allows for iterations and adaptive solutions [27], [28]. For the capstone class, the project team consists of the analysts (the learners), a process expert which serves as a coach and facilitator (the instructor), and the product owners who represent the stakeholders (the sponsors) [27]. Once the project plan is defined, the project is broken down into 3-week long sprints where learners define short-term sprint goals, evaluate the sprint results, and then plan the next sprint [28]. The idea is that the project plan will serve as the product backlog to give a broad overview of the project and give a starting point to plan the sprints which will be conducted in the second

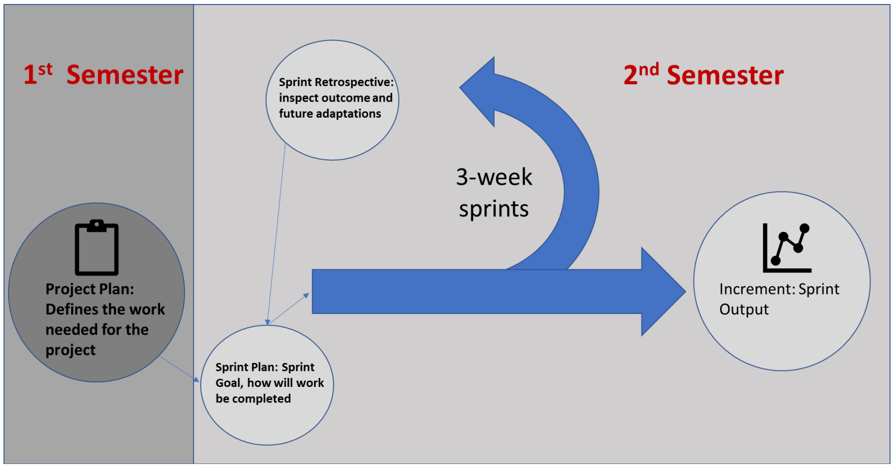

Figure 5: Breakdown of first and second semester capstone courses using an Agile framework.

semester. Fig. 5 shows a schematic breakdown of the first and second semester capstone courses.
Using an Agile project management structure provides a means to define the project and begin to build the data story. Defining the project plan gives focus to the project and enables learners to begin defining the important elements of their data story. This clarifies what information is important to communicate, but also, what information to omit [13], [21].

The first semester concludes with an initial project sprint so that the learners proceed past the first two stages of the CRISP-DM workflow and engage with Data Preparation, Modeling, and Evaluation of results for their project. The learners also provide a mid-semester and final presentation to communicate project results.

The second semester is divided into three separate sprints which are each three weeks in length. Each sprint begins with a sprint planning exercise where the teams use the project plan generated in the first semester and create a more refined plan for the individual sprint. The teams generate a sprint report which is the data story, or "increment" (output), of the sprint. It is expected that the teams will learn more about the project as they progress through the sprints which may require deviations or modifications from the original project plan. Learners are encouraged to modify the project plan as they progress through the project, using the Agile methodology to explain pivots necessary to keep the project on track for the sponsors.

4. Reflections and Recommendations

There are numerous challenges associated with designing a graduate curriculum for working professionals to adopt data science methods. Graduate programs are much shorter than undergraduate programs, and one challenge is that rather than taking a deeper dive into a topic from undergraduate programming, this graduate program is designed to pivot learners to new data science methods. One challenge is that despite many learners being new to data and computer science, these methods must be presented in a way that is accessible to novice computer science learners in only a few courses. The interdisciplinarity of the program also means that there are only a handful of courses that present computer science topics and gives learners the opportunity to apply these methods in practice.

The learners are also very diverse, coming from different educational backgrounds (Fig. 1), ages ranging from mid-twenties to early fifties, and various professions. Since the program is a graduate program, supplemental programming in addition to formal instruction is being explored to provide additional support to learners. The program has recruited data analytics undergraduate students employed at a university research institute to tutor these learners in basic programming skills. Over 70 tutoring sessions were scheduled in the first semester the tutoring program was offered, with tutoring session in R, Python, and statistical concepts. Most learners choosing to participate in tutoring engaged multiple times throughout the semesters. All learners who have engaged with the tutoring program successfully completed the formal classes they were enrolled in for that semester.

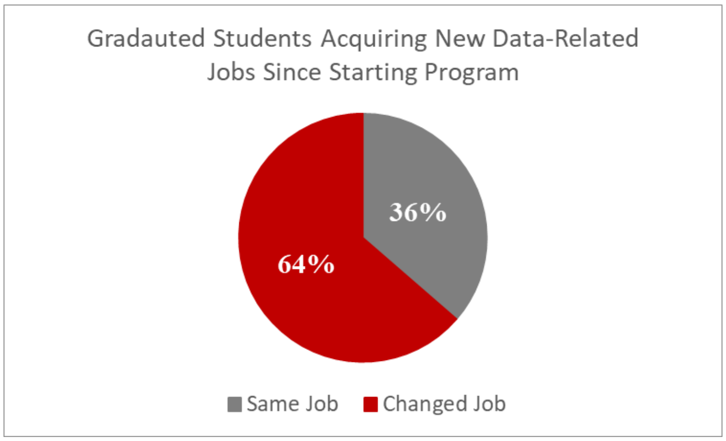

Figure 6: Graduates who have acquired new data-related positions since starting the program.

The goal of this program is to transition working professionals from various backgrounds to data focused careers. To date, two cohorts have graduated from the program with 19 learners and 14 learners respectively in each cohort. An analysis of LinkedIn profiles indicates that 64% of graduated learners have secured new positions in data related careers since starting the program. Fig. 6 shows the breakdown of employment status after graduation for the total number of graduated learners.

The program also created an online, asynchronous orientation course to provide information and resources during the summer prior to the start of the first semester. The orientation course reviews logistical items related to graduate programming, and also provides a broad overview of the interdisciplinary areas: computer science, statistics, and design and how they work together in the overarching curriculum. In addition, relevant math concepts are presented so that the students who may be experiencing math anxiety can work during the summer to feel more confident at the beginning of the program. Adult learners may have lower math self-efficacy which may be due to several factors including time away from the classroom, complicated life situations, and new modalities in learning that were not common during their prior educational experiences [12]. Additionally, many of these learners may not have an extensive math background and could be feeling apprehensive about the analytical nature of the program. Providing a math refresher can be a resource for learners seeking to alleviate some of the stress related to beginning a graduate program prior to the start of the program. The intention is to ease these learners into the program, and improve the learner experience, particularly at the start of the program. The orientation course also provides some resources on coding such as MOOC's which the learners, as well as some tips on approaches to troubleshooting code.

4.1 Reflection for Foundational Courses

The program is designed for, and attracts, a diverse population with respect to computer-science education and experience. While this can create challenges for all participants, the typical learner has strong expertise in their chosen field, strong work ethic, and considerable maturity as adult learners which serves them well in graduate programming [29], [30]. This leads to productive discussions of practical applications of data analytics, learners helping other learners, and unexpected insights.

The expectation that a learner who has never written code in any programming language can learn Python in one semester is a high bar. For example, a learner might feel that code that is "almost" right (e.g., using a semi-colon instead of a colon in a Python "if" statement) should "almost" work. This can lead to frustration.

Given this frustration, it is important to remind learners often that the instructors and staff are available to help: "If you spend more than a half hour trying to get this function to work, post a question to the Teams chat and we will help." From our perspective, it is more important for a learner to make an honest attempt to complete an assignment and then ask for help than it is for the learner to go it alone.

It also is important to provide guidance on how to write and debug code effectively. For example, when writing SQL, don't write a long, multi-line statement and expect it to work. Write a very simple SQL statement and make it work. Then incrementally add clauses and make them work. These practices are second nature to experienced programmers, but new to some learners.

On the other end of the spectrum, some learners are already quite skilled at Python programming and other topics covered by the curriculum. The goal in these cases is to provide challenging supplemental components in the assignments to keep them engaged in the coursework.

Regarding the overall computer-science curriculum, we try to balance practical programming skills with deeper understanding of the fundamental concepts. As mentioned above, some learners become frustrated when they are doing non-programming assignments – wanting to get back to the "real" work. We continually point back to the basic goal of data analytics: problem solving.

Finally, we cast the curriculum as an entry point into life-long learning in computer science and point to other resources and topics that are valuable to explore.

4.2 Reflection for Data Learning and Mining Course

Practical Learning and Mining for Big Data is co-taught between the Department of Statistics and the Department of Computer Science and Engineering; this calls for careful preparation and delegation of tasks well before the course begins. However, students have commented that they appreciate the distinct perspective from both faculty, so it is a worthwhile challenge. The wide range of learner backgrounds and perspectives should be used as an advantage in a course such as this [29], [30]. Data sets relevant to students' professional backgrounds can motivate the content, demonstrate application in professional contexts, and give students experience summarizing findings in an accessible way. The focus on technical writing to explain and summarize findings prepares students for the more formal challenge that awaits them in the program capstone courses. Feedback should be given regularly regarding both the technical aspects of the course as well as their written work.

4.3 Reflection for 2- Semester Capstone Sequence

The students in each of the two offerings of the capstone course sequence were eager to begin the capstone course. However, they quickly became frustrated with the focus on developing the business understanding of the project and many learners indicated that they wanted to jump into an analysis of the data. Learners described feeling apprehensive of losing practice with their recently acquired computer science skills since they were not actively coding from the start of the project. The second offering of the course reorganized the first semester course content to incorporate more opportunities for analysis, but the focus on Lane 4, Data Storytelling, remained to emphasize the importance of the early stages in the project. Without clearly defined objectives and goals, telling a clear story with data becomes an almost impossible task, regardless of the learners' desire to jump into an analysis with the data. The unique 2-semester structure of the capstone experience allows for time to be spent on this critical early stage of the project.

Overall, the learners completed the assignments, and the projects were successful. Two sponsors have participated in all offerings of the program's capstone, and others have stated that they would participate again indicating a positive sponsor experience with the program. Shortcomings in the assignments were primarily in the interpretation and presentation of results. This included improper formatting, such as missing Fig. labels and difficult to read visualizations, documents which fall short of expectations for writing quality and organization, and statements indicating "good" performance of data models but lacked definitions for "good performance" and did not relate the results to the larger project context.

Learners were also uncomfortable with the ambiguous nature of these projects, although a few stated that their favorite component of the class was the freedom they had to explore the project. Often, results were inconclusive or did not produce expected answers. Despite assurances that unexpected results from an analysis, including analyses that did not work, can lead to insight about a problem, a few learners felt as if the project had failed in some way because they were unable to come up with a single conclusion or answer.

5. Conclusions

Creating an interdisciplinary graduate program to pivot working professionals to data science careers is a compelling challenge. Although more than half of graduates have made successful transitions to data focused careers, opportunities remain to refine the curriculum and improve learner outcomes. The program will continue to explore providing supplementary programming to learners such as the tutoring program. This includes providing resources prior to the start of the program so that learners can ease into these new concepts and avoid being overwhelmed from the start of the program.

Although the program provides coursework in design and incorporates design concepts in all classes, stronger connections between these disciplines in needed for the learners to emphasize the importance of data storytelling. This is an ongoing challenge that needs to be explored in more depth.

In addition to providing supplementary programming, the curriculum will continue to be fine-tuned to ensure that learner outcomes are being addressed. This includes focusing on primary topics and recognizing that the curriculum is short and is intended as a starting point for continuous lifelong learning. Data science and analytics is a constantly changing field. Providing individuals skills to continue to learn new methods in data science and analytics is critical to the development of a vibrant and sustainable data centered workforce.

References

[1] M. Ghasemaghaei and G. Calic, "Does big data enhance firm innovation competency? The mediating role of data-driven insights," *Journal of Business Research*, vol. 104, pp. 69–84, Nov. 2019, doi: 10.1016/j.jbusres.2019.07.006.

[2] M. Younas, "Research challenges of big data," *Service Oriented Computing and Applications*, vol. 13, no. 2, pp. 105–107, 2019, doi: 10.1007/s11761-019-00265-x.

[3] M. Naeem *et. al.*, "Trends and future perspective challenges in big data," in *Advances in Intelligent Data Analysis and Applications: Proceedings of the Sixth Euro-China Conference on Intelligent Data Analysis and Applications*, *15-18 October 2019 Arad, Romania.* Springer Singapore, 2022, pp. 309–325.

[4] Bureau of Labor Statistics, "Data Scientists," *Occupational Outlook Handbook*. 2023. Accessed: Aug. 15, 2023. [Online]. Available: https://www.bls.gov/ooh/math/data-scientists.htm

[5] J. A. Rosendale, "Gauging the value of MOOCs," *Higher Education, Skills and Work-Based Learning*, vol. 7, no. 2, pp. 141–154, Jan. 2017, doi: 10.1108/HESWBL-09-2016-0065.

[6] National Academies of Sciences, "Data science for undergraduates: Opportunities and options," Washington, DC, 2018.

[7] L. Modenos, "No, nontraditional Is not the new traditional," *Adult Learning*, vol. 31, no. 3, pp. 134–136, Jul. 2020, doi: 10.1177/1045159520941082.

[8] Fortune, "Best online master's in data science programs in 2023," *Fortune*, 2023.

[9] R. S. Malik, "90 MS data science programs in the USA - Complete List," Medium. Accessed: Jul. 28, 2023. [Online]. Available: https://medium.com/@rijulsinghmalik/90-ms-data-science-programs-in-the-usa-complete-list-4a8ff82443b7

[10] T. Heaney, "Adult education for social change: From center stage to the wings and back again," Information Series No.365, pp.1-29, 1996

[11] C. de Brey *et. al.*, "Status and trends in the education of racial and ethnic groups 2018. NCES 2019-038" *National Center for Education Statistics,* 2019.

[12] M. E. Johnson, A. Albizri, and R. Jain, “Exploratory analysis to identify concepts, skills, knowledge, and tools to educate business analytics practitioners,” *Decision Sciences Journal of Innovative Education*, vol. 18, no. 1, pp. 90–118, Jan. 2020, doi: 10.1111/dsji.12195.

[13] V. Echeverria *et al.*, “Driving data storytelling from learning design,” in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, in LAK 2018. New York, NY, USA: Association for Computing Machinery, 2018, pp. 131–140. doi: 10.1145/3170358.3170380.

[14] H. Karbasian and A. Johri, “Insights for curriculum development: Identifying emerging data science topics through analysis of Q&A communities,” in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, in SIGCSE 2020. New York, NY, USA: Association for Computing Machinery, 2020, pp. 192–198. doi: 10.1145/3328778.3366817.

[15] ACM and IEEE Computer Society, *Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science*. New York, NY, USA: Association for Computing Machinery, 2013.

[16] “Ohio Supercomputer Center.” Accessed: Jan. 16, 2019. [Online]. Available: https://www.osc.edu/

[17] J. S. Saltz, “CRISP-DM for data science: Strengths, weaknesses and potential next steps,” in *Proceedings - 2021 IEEE International Conference on Big Data, Big Data 2021*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 2337–2344. doi: 10.1109/BigData52589.2021.9671634.

[18] J. Saltz, “CRISP-DM is still the most popular framework for executing data science projects,” Data Science Process Alliance, 2022.

[19] N. Hotz, “What is CRISP-DM,” Data Science Process Alliance. Accessed: Aug. 15, 2023. [Online]. Available: https://www.datascience-pm.com/crisp-dm-2/

[20] M. Kelly, R. Longjohn, and K. Nottingham, “The UCI Machine Learning Repository.” Accessed: Aug. 15, 2023. [Online]. Available: https://archive.ics.uci.edu

[21] B. Dykes, *Effective data storytelling: how to drive change with data, narrative, and visuals.* John Wiley & Sons, 2019.

[22] P. Chapman *et al.*, “CRISP-DM 1.0,” in *4th CRISP-DM SIG Workshop* , Brussels, Mar. 1999.

[23] D. Spradlin, “Are you solving the right problem?,” *Harvard Busines Review*, pp. 85–94, Sep. 2012.

[24] E. A. Vance *et. al.*, “Asking great questions,” *Stat*, vol. 11, no. 1, Dec. 2022, doi: 10.1002/sta4.471.

[25] R. Lingard and S. Barkataki, “Teaching teamwork in engineering and computer science,” in *2011 Frontiers in Education Conference (FIE)*, 2011, pp. F1C-1-F1C-5. doi: 10.1109/FIE.2011.6143000.

[26] L. Gueldenzoph Snyder, “Teaching teams about teamwork: Preparation, practice, and performance review,” *Business Communication Quarterly*, vol. 72, no. 1, pp. 74–79, 2009.

[27] J. Saltz, “Data driven scrum,” *Data Science Process Alliance*. May 30, 2022. Accessed: Aug. 15, 2023. [Online]. Available: https://www.datascience-pm.com/data-driven-scrum/

[28] K. Schwaber and J. Sutherland, “The 2020 Scrum Guide,” 2020.

[29] R. J. Wlodkowski, *Enhancing adult motivation to learn : a comprehensive guide for teaching all adults*. San Francisco: Jossey-Bass, A Wiley Imprint, 2008.

[30] M. S. Knowles, *The modern practice of adult education : From pedagogy to andragogy*. [Wilton, Conn.]; Chicago: Association Press ; Follett Pub. Co., 1980.