Statement of Teaching Philosophy

Daryl DeFord

As a graduate student, postdoc, and fauclty member I have had opportunities to teach a wide variety of courses in mathematics, statistics, and data analytics at both the undergraduate and graduate levels, develop new courses of my own, refine previously taught courses, and explore unique pedagogical opportunities outside the formal classroom. I use the following three guiding principles for evaluating and determining my actions as an instructor, from smaller choices, such as the inclusion of a particular example, to large-scale design of the course. I believe that my decisions should support students developing their critical thinking skills with mathematics, improve their exposure to formal mathematical reasoning and intuition, and challenge every student to deepen their understanding of mathematics. In addition to these philosophical considerations I am particularly mindful of my role as an instructor in preparing students for their next steps once they leave the classroom. This takes different forms depending on the course context, as undergraduate data analytics majors have very different needs than graduate students in pure mathematics or statistics, but it informs many of the specific details that I describe below.

For data-centric courses my goals include helping the students engage with a very broad view of data and all of the places that it can impact their lives. By incorporating readings and discussions groups, I encourage them to view themselves as active participants rather than just passive consumers of data analysis. This is also reinforced through the assessment design in these courses, where projects encourage them to pursue topics they are passionate about and to engage in end-to-end data work. In addition to traditional learning outcomes from my courses, I also include motivating questions in the syllabus to help make more concrete the learning achievements the students are making throughout the semester. My goal is for students to leave these courses with clear benchmarks demonstrating what they have accomplished and for them to be immediately ready to translate the skills that they have gained to their chosen career paths. My focus on this viewpoint has been significantly informed and supported by my participation in the Core-to-Career pedagogy development program at WSU.

I strive to maintain an engaging and inclusive classroom environment, preferring discussions and interactive experiences to passive lecturing. This is particularly true for courses with a programming component, where I often devote significant effort to creating interactive notebooks and code examples that the students can modify to better explore the key topics. Even in more theoretical courses, I tend to use motivating problems or challenges that the students work on in groups to introduce new concepts and ideas by exploration. In graduate courses this often takes the form of a warm up problem to be completed together at the beginning of class or a 'problem of the day' to be discussed before the next meeting. These are often difficult challenges that the students are not expected to solve completely but rather opportunities to struggle and experiment in a friendly and supportive environment, as detailed in my guiding principles.

Beyond classroom instruction, I also currently serve as the research advisor for five PhD students and five MS students. Independently and through the STEM Research and Mentoring Program I also supervise several undergraduate researchers. I often incorporate examples from my research in teaching, including offering graduate topics courses on state-of-the-art methods and incorporating real-world data examples in undergraduate courses as discussed below. In order to develop my teaching and course design skills I participated in the WSU Core-to-Career program with a focus on revising and improving DATA 115. I am also actively engaged with a community of educators interested in modernizing linear algebra education and was recently invited to give a talk at a special session of a SIAM conference by the organizers of the International Linear Algebra Society education group as a result of a podcast about the interaction between network science and linear algebra that I recorded for their pedagogical website. In 2021 I helped teach the JMM Short Course on complex networks and in 2023 I was invited to participate in the AMS Engaged Pedagogy Series, developing materials and presentations to train other lecturers in adding coursework components on computational redistricting.

In order to highlight my approach to course design and instruction I will give several examples from my development of Data 115: Introduction to Data Analytics at WSU. This is a foundational and introductory course for students in the Data Analytics program but it also has become a popular course for advanced students from a variety of disciplines who want to develop their data skills before going on the job market.

In recent semesters a plurality of students have been seniors, which is quite unusual for a 100-level course. This creates some unique challenges for the course because the student backgrounds are so diverse - most of the first year Data students have no programming experience at all, while the advanced students often have taken full sequences from the computer science department. In my design of the course, I attempted to treat this as an opportunity, leveraging the advanced students' experiences to help guide the students just starting out on their data pathways.

Project-based learning allows students to engage with the material over a longer period of time, and also provides opportunities for them to develop skills that would be difficult to incorporate on a single weekly assignment. For Data 115 I introduced two projects for the students - an individual Personal Dataset Project (PDP) and a final group dataset analysis. These projects are complementary, allowing students to focus on different skill sets and applications of data methods throughout the semester. The PDP focuses on building and cleaning a dataset related to each student's interest, while the final project focuses on performing analytic techniques on a published dataset. Students get an opportunity to practice their presentation skills for both projects and also reinforce the course focus on visualization design and professional writing.

The initial motivation for the PDP was to provide students with the opportunity to leave the course with a dataset that they are passionate about and understand deeply. They are then able to use this dataset as a baseline test case when they encounter descriptions of methods in later courses in the major but aren't provided with relevant data. This has been a common concern from students who sometimes struggle when presented with abstract methods without examples in their later coursework. Additionally, the final product is a polished GitHub repository that the students can incorporate in their job market portfolios. In order to provide scaffolding for the project I broke the final components into a sequence of individual steps that are included in every other weekly homework assignment, starting at the beginning of the semester. This helps students make consistent progress throughout the semester and by tying the individual steps to what we are covering that week it also helps them see how the course material is relevant to real data analysis projects. Finally, going through all of the stages of data collection and cleaning is a valuable exercise as most data professionals estimate that the majority of their time is spent on similar tasks.

The reading discussion component of the class provides students an opportunity to engage with modern data analytics work with a goal towards developing critical thinking skills. In this case developing the ability to be actively critical of published work is incredibly important, as we frequently encounter examples where the presented analysis or visualizations are not accurate or efficient representations of the underlying data. Each week I provide the students with a couple of published data analyses on a topic relevant to that week's course material and a set of discussion questions that I've generated to help them engage with the material. We do the first set together in the first week so I can explain the intent of the questions with an example about agricultural production in Washington state, motivated by signs on the highways leading here that say "Welcome to Whitman County - The nation's leading wheat producing county." This leads to a very natural discussion about how the choice of metric impacts the results that are presented and how informed analysis can help students better understand the world around them. In subsequent weeks I split them into smaller groups to have these discussions and to write up their responses to the questions. The student engagement and responses have been consistently impressive and I have found facilitating these conversations to be one of the most exciting teaching experiences in the course.

Another key feature of the reading discussions that I have tried to reinforce after teaching the course for the first time was the idea of of helping students view themselves as 'data people.' This is particularly important for students who had previous work experience around data but not with analytics specifically and students who have not had good experiences in previous quantitative courses. In my discussions with students, particularly non-traditional students on the Vancouver and Everett campuses, I learned a lot about ways to help them view their prior experiences as relevant to modern data work. A particularly impactful conversation with a woman who had previously worked for a school district doing data entry reinforced this for me as a critical function of this course. Discussing ways for her to translate her experiences to the context of the class and data work more broadly transformed her progress through the course and I've tried to have as many similar conversations as possible with students individually. To integrate

these conversations into the reading discussions I adjusted the questions, particularly early in the semester, to focus on brainstorming how their prior knowledge can contribute to the analysis.

One of my favorite assessments that I created for the class challenges the students to 'clean up' a messy dataset to perform a simple calculation. After describing internal consistency checks and walking the students through several examples of column-based operations I provide them with a synthetic dataset based on my research work in redistricting and ask them to determine what appears to be a simple ratio of the column sums. However, inside the dataset I have included various distortions including duplicate entries, duplicate ids with inequivalent values, negative values, rows with incomplete data, and rows with inconsistent or impossible values, motivated by real data. Students must detect, diagnose, and decide how to address these issues to get a reasonable estimate for their final answer. This leads to a natural class discussion about the unreliability of real data sets and I walk them through the real-world example (complete with negative vote totals in the official state dataset and the FOIA request that I filed over it) to show the relevance of the exercise.

It is common that no two students get the same answer to this problem, which highlights one of the key questions that I encourage students to engage with throughout the course - What do we do when there is no 'right' answer? This is a fundamental part of modern data practice but something that often gets left out of introductory courses, where all of the presented datasets are already cleaned and verified. In 115 we start in week 1 by discussing the realities of actual data work and focus on what I call a 'Document and Justify' or 'ReCheck and Report' approach to empirical work in the face of uncertainty throughout the course. This is also supported by the discussion questions and readings where students are asked to engage with this concern in the published work of experts in the field.

Another one of my main topics of emphasis in data course development has been the incorporation of data ethics modules. To me this is a fundamental part of data education, as we are expecting our students to be informed and responsible scientists. As an introductory course, this module focuses on a professional ethics perspective and an introduction to common case studies in algorithmic bias and dataset anonymity. These materials build on my experience teaching research ethics for three years as a graduate student. I further developed some of these materials into a presentation to coincide with WSU's Common Read last year. For more advanced data courses I discuss techniques for analyzing and mitigating ethical issues in data work and recently designed a new graduate course on Responsible Data Science that focuses entirely on these issues.

The materials that I prepared for day-to-day class meetings consisted of slides that I provided for the students and Jupyter Notebooks or RMarkdown files with live examples that I would work through in class with the students. These were set up so that students could follow along during lecture and experiment with running the methods and changing the values to build up intuition for the methods. This also helped them with the transition to programming and was incorporated into the weekly assignments by providing them with a starting point for the code that they needed. I also constructed several guides to various computational aspects of the class for the students to review. This is a part of my teaching approach, developing helpful guides to help students focus on the important material, rather than memorization. This also matches the students eventual working environments after graduation, where knowing how to search and use internal documentation is an important skill.

While the specific material varies between classes and disciplines, I try to take a similarly expansive approach in all of my pedagogical activities. I hope the examples above demonstrate that I have devoted significant effort as a faculty member to developing my skills as an educator and have tried to thoughtfully apply my perspectives to designing learning experiences that will benefit students both inside and outside the classroom. This includes outreach efforts and other teaching that I do outside the formal classroom. For example, the materials I presented in the AMS Engaged Pedagogy Program were originally developed for the research fellows I mentored through the Data Science for Social Good program, where my efforts were widely regarded by the program staff and data scientists.

I am constantly working to develop my skills as an instructor and I view each interaction with students, whether in formal class meetings, office hours, or informal discussion, as an opportunity to put my guiding principles into practice for the benefit of my students.