UNIVERSITY OF SAN FRANCISCO | College of Arts and Sciences

BACHELOR OF SCIENCE IN DATA SCIENCE

ASSESSMENT REPORT
ACADEMIC YEAR 2022 – 2023

I. LOGISTICS & PROGRAM LEARNING OUTCOMES

1. **Please indicate the name and email of the program contact person to whom feedback should be sent (usually Chair, Program Director, or Faculty Assessment Coordinator).**

   Program Director: Daniel O'Connor (doconnor@usfca.edu)

2. **Were any changes made to the program mission statement since the last assessment cycle in October 2022? Kindly state "Yes" or "No." Please provide the current mission statement below. If you are submitting an aggregate report, please provide the current mission statements of both the major and the minor program.**

   No

   To deliver a high-quality data science program that instructs students in the theory and practice of mathematical and computational analysis of applied data driven problems, and to graduate students with appropriate experience in industry-standard data science tools.

3. **Were any changes made to the program learning outcomes (PLOs) since the last assessment cycle in October 2022? Kindly state "Yes" or "No." Please provide the current PLOs below. If you are submitting an aggregate report, please provide the current PLOs for both the major and the minor programs.**

   No

   - [PLO1] Analyze information critically and logically in a mathematical setting.
   - [PLO2] Reformulate and solve problems in an abstract framework.
   - [PLO3] Express mathematical results verbally, working individually and in collaborative groups.
   - [PLO4] Apply mathematical techniques to specific problem domains
   - [PLO5] Demonstrate competence with programming concepts, including software development techniques and data structures
   - [PLO6] Apply mathematical and computational techniques to real-world problems involving large, complex data sets.
   - [PLO7] Visualize, present and communicate analytical results.

4. **Which particular Program Learning Outcome(s) did you assess for the academic year 2022-2023?**

   PLO1, PLO4, PLO5, PLO7

II. METHODOLOGY

5. **Describe the methodology that you used to assess the PLO(s).**

   For example, "the department used questions that were inputted in the final examination pertaining directly to the <said PLO>. An independent group of faculty (not teaching the course) then evaluated the responses to the questions and gave the students a grade for responses to those questions."

   We directly assessed all graduating seniors with an end-of-degree exam given in the Spring 2023 semester. This exam consisted of 14 multiple choice questions spanning topics from the required curriculum. This exact exam, with the same 14 questions, was also given to previous cohorts of graduating seniors in the Spring 2022, Spring 2021, Spring 2019, Spring 2018, and Spring 2017 semesters. Note that this exit exam was not given during the Spring 2020 semester due to the pandemic. It is our intention to continually assess our students and, by extension, the program by annually giving the graduating seniors the same exam. This will provide objective and comparable year-over-year data with which we can evaluate the effectiveness of the program. We presently have six years of data whose results and findings I will discuss in the next section. The exit exam is attached as a separate document.

This year, we also assessed PLO7 by requiring MATH 373 students to create powerpoint slides which explain and summarize what they did and what results they obtained for their final projects. The final project involved training a neural network (U-Net architecture) to label nuclei in cell microscopy images. These slide decks were judged on the following criteria:

- **Clarity**. Are the slides easy to understand? Are the explanations correct, concise, and illuminating?
- **Level of technical detail**. Are the methods and results explained in sufficient detail?
- **Storytelling**. Do the slides tell an engaging story?

Each slide deck received a rating of either "poor", "satisfactory", "good", or "excellent" for each of these three criteria.

III. RESULTS & MAJOR FINDINGS

6. **What are the major takeaways from your assessment exercise?**

This section is for you to highlight the results of the exercise. Pertinent information here would include:

Exit exam results from the years 2017 – 2019, 2021 - 2023 are summarized in Figure 1.

Graduating seniors who took the assessment test performed noticeably better on the exam in 2023 than in previous years. Some possible explanations for the jump in test scores will be discussed below. The mean test score from 2023 was 10.17, whereas the mean test score from 2022 was 6.33. A two-sample t-test of the hypothesis that the difference in mean scores for the years 2023 and 2022 is not zero yields a p-value of .0014, indicating that the difference in mean scores is statistically significant.

The exit exam scores in 2023 were stronger overall than in all previous years. A two-sample t-test of the hypothesis that the mean score from 2023 is greater than the mean of the exit exam scores from the years 2017-2019, 2021-2023 yields a p-value of .0034, indicating that the difference in mean scores is statistically significant.

However, there is an important caveat – the exit exams were taken remotely and self-proctored in 2023. This likely contributed to the boost in 2023 test scores. We should also note that the 2023 sample size n = 6 is fairly small, and several of our top data science majors were included in this sample. Another possible contributing factor for the boost in test scores is that, after a period of high faculty turnover, our data science professors have gained experience. Moreover, our most experienced data science professors James Wilson and Steve Devlin have both returned to the USF Math & Statistics department. They taught key courses for this cohort.

In Table 1 we define four levels of mastery, map those to ranges of test scores and identify the percentage of students achieving each level. Table 1 includes all test scores

from the years 2017-2019, 2021-2023. This information is also depicted in Figure 2. In summary, 70% of the students scored either satisfactory, good, or excellent. It's an important goal to find ways to bring these scores up.

Level of Mastery	Exam Scores	No. Students	% Students
Poor	0-6	15	30%
Satisfactory	7-9	22	44%
Good	10-11	11	22%
Excellent	12-14	2	4%

Table 1: BSDS Levels of Mastery (2017 – 2019, 2021-2023 aggregate)

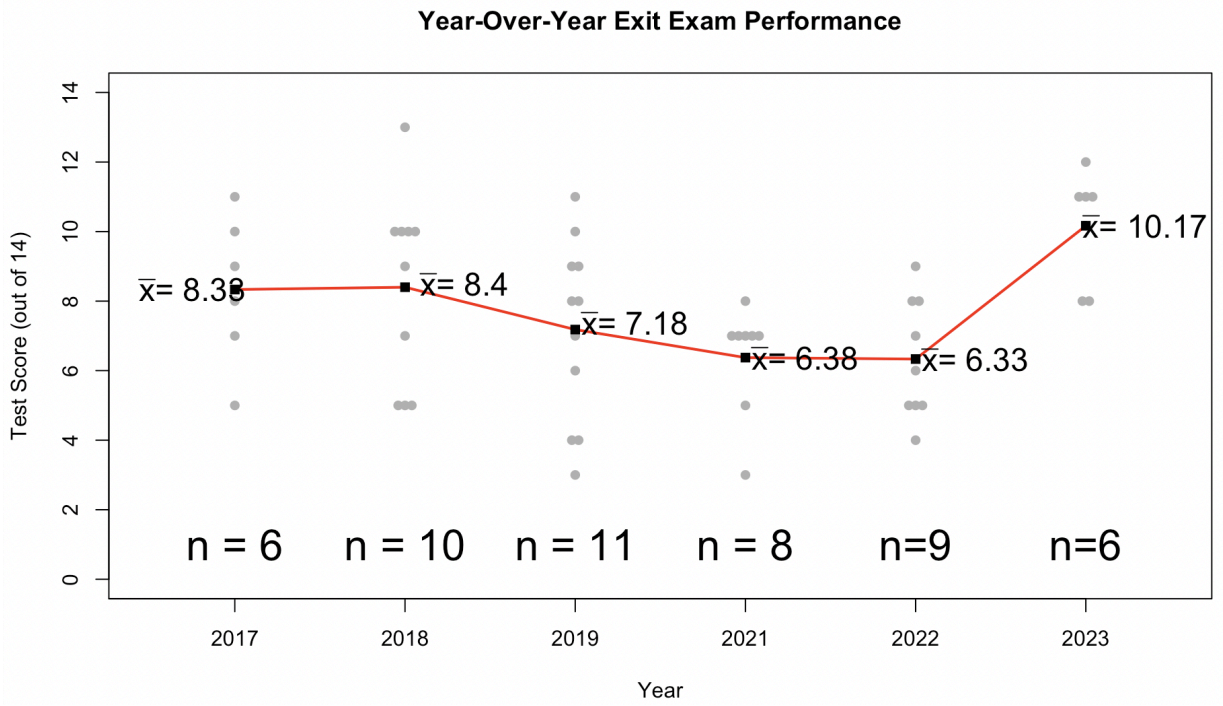

Figure 1: BSDS Exit Exam Scores Year over Year

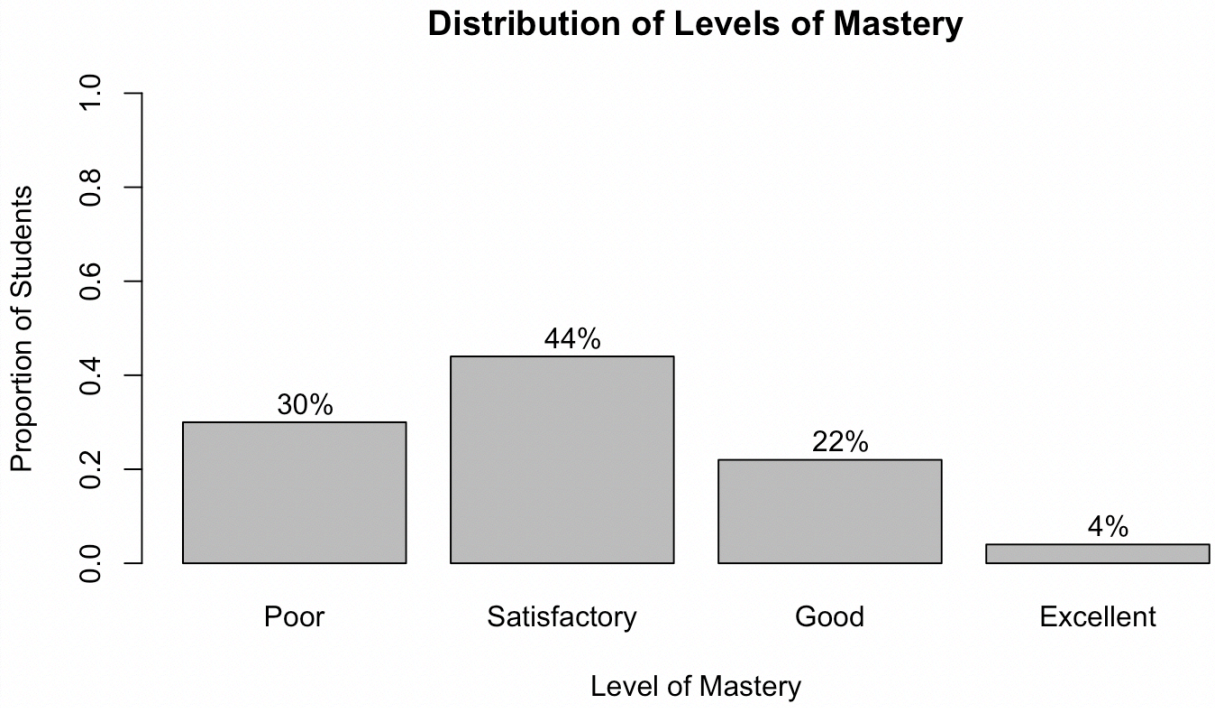

Figure 2: BSDS Levels of Mastery (2017 – 2019, 2021-2023 aggregate)

As mentioned in section II, in the Spring 2023 semester we assessed PLO7 by rating slide decks for MATH 373 final projects according to three criteria: clarity, level of technical detail, and storytelling. These ratings are summarized in figures 4, 5, and 6. An example slide is shown in figure 3.

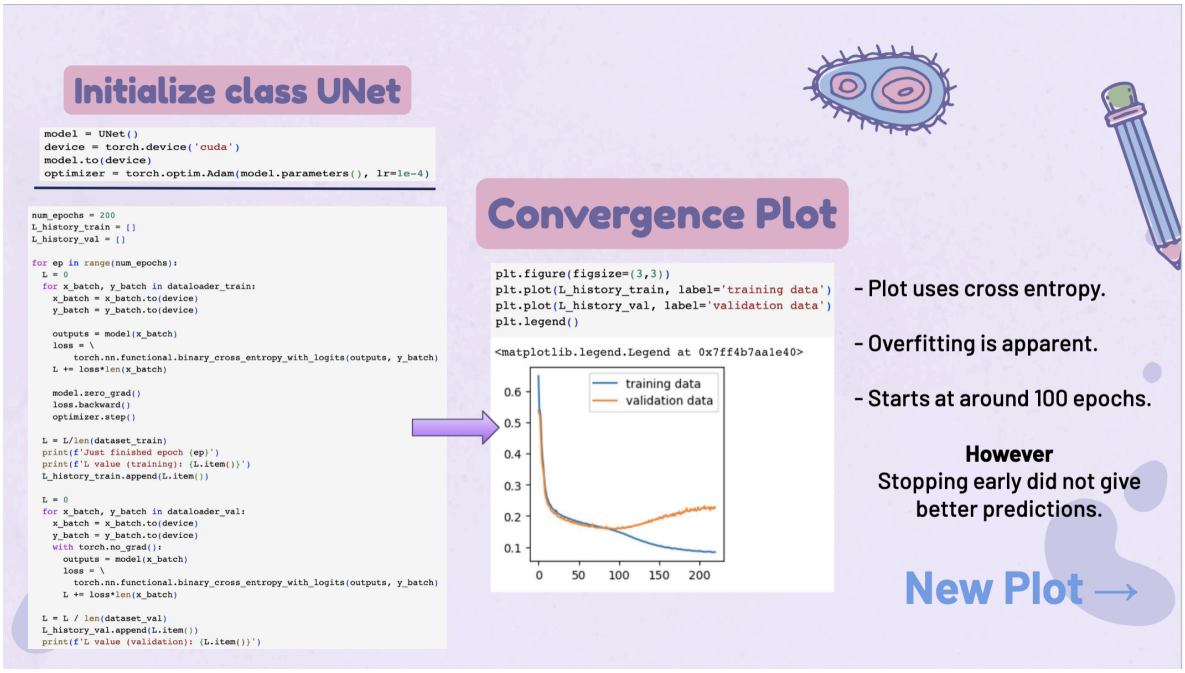

Figure 3: Example slide from a MATH 373 final project (Spring 2023).

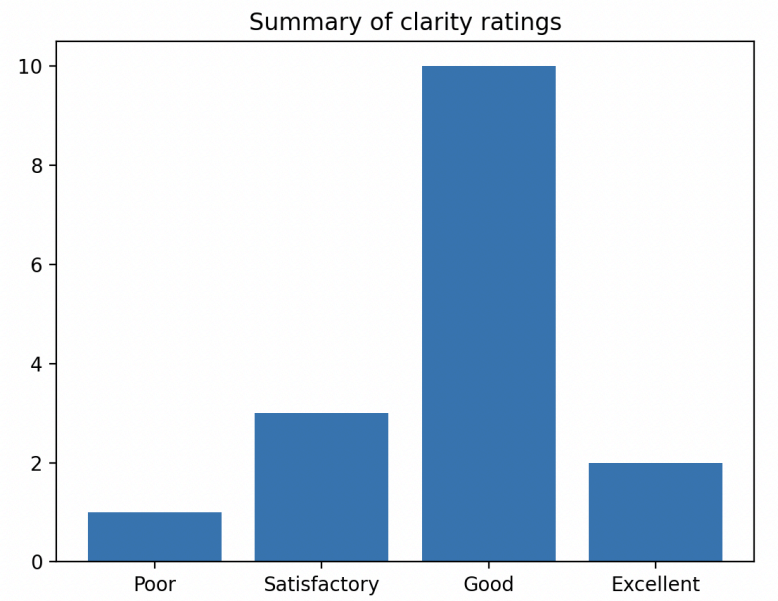

Figure 4: Summary of clarity ratings for MATH 373 final project slide decks (Spring 2023).

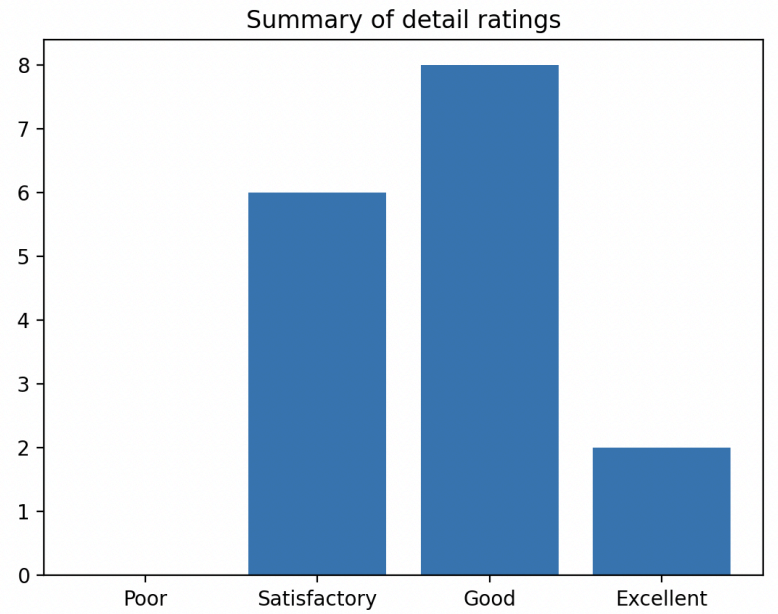

Figure 5: Summary of ratings for "level of technical detail" for MATH 373 final project slide decks (Spring 2023).

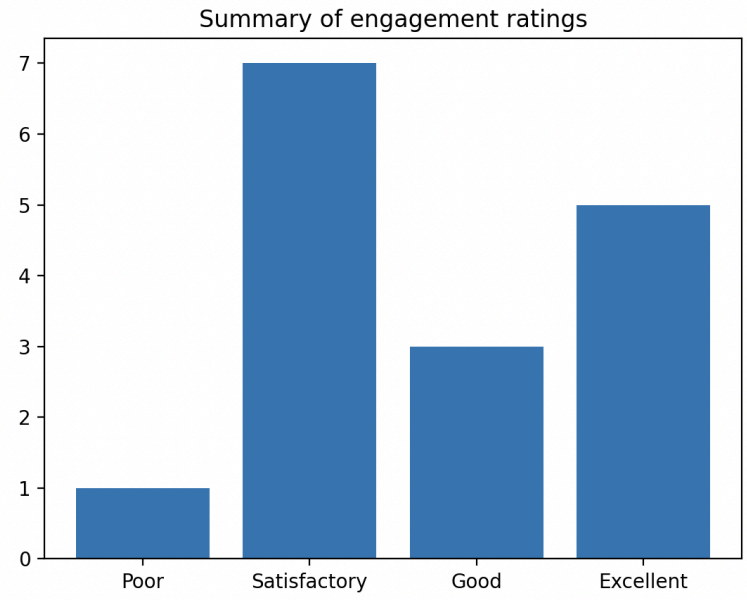

Figure 6: Summary of storytelling ratings for MATH 373 final project slide decks (Spring 2023).

IV. CLOSING THE LOOP

7. **Based on your results, what changes/modifications are you planning in order to achieve the desired level of mastery in the assessed learning outcome? This section could also address more long-term planning that your department/program is considering and does not require that any changes need to be implemented in the next academic year itself.**

The students have most often struggled with questions 1, 11, 12 and 13, which are technical questions about probability, statistics, and linear algebra. Students have also had trouble with problems 3, 6, and 8, which indicates a struggle with PLO4 (applying mathematical techniques to specific problem domains). The topics being tested by these questions are:

- Conditional probability

- Eigenvalue calculation
- Least squares
- Likelihood estimation

All of these topics are tied to specific classes (MATH 230, MATH 370, MATH 371, MATH 372). Based on these findings we plan to ensure these topics (in these classes) are clearly emphasized and that their importance beyond the classroom is highlighted. This should help to improve student performance on these questions and the level of mastery associated with this learning outcome. We have recently hired a new term professor Sam Roven to help with teaching linear algebra (among other classes), and it's very helpful that James Wilson is back at USF, now solely teaching undergraduate data science majors. James has been on sabbatical this year, but he will be back in full force next year.

8. **What were the most important suggestions/feedback from the FDCD on your last assessment report (for academic year 2021-2022, submitted in October 2022)? How did you incorporate or address the suggestion(s) in this report?**

We always appreciate the insightful & thoughtful feedback. Based on last year's feedback, we added an assessment of PLO7, requiring MATH 373 students to submit slide decks reporting on their final projects. The slides were rated based on clarity, level of technical detail, and storytelling.

V. ADDITIONAL MATERIALS

Below we include the following additional materials:

- End of the degree exam
- Curricular maps