SAINT LOUIS UNIVERSITY.

Program-Level Assessment: Annual Report

Program Name (no acronyms): Data Science	Department: Math/Stats and Computer Science
Degree or Certificate Level: BS	College/School: Arts & Sciences
Date (Month/Year): September/2021	Assessment Contact: Darrin Speegle

In what year was the data upon which this report is based collected? AY 2020-2021

In what year was the program's assessment plan most recently reviewed/updated? AY 2019-2020

1. Student Learning Outcomes

Which of the program's student learning outcomes were assessed in this annual assessment cycle? (Please list the full, complete learning outcome statements and not just numbers, e.g., Outcomes 1 and 2.)

1. Graduates will be able to visualize data to facilitate the effective presentation of data-driven insights.
2. Graduates will apply statistics to analyze data sets.
3. Graduates will be able to use technical skills in predictive modeling.

2. Assessment Methods: Artifacts of Student Learning

Which artifacts of student learning were used to determine if students achieved the outcome(s)? Please describe and identify the course(s) in which these artifacts were collected. Clarify if any such courses were offered a) online, b) at the Madrid campus, or c) at any other off-campus location.

Five sections of four courses were used for assessing student learning outcomes: DATA 1800, DATA 2800, STAT 3850 (2 sections) and STAT 4840. In STAT 3850 and STAT 4840, instructors chose course learning outcomes that related to the program learning outcomes described in part 1 and asked a question on the final exam relating to the chosen course learning outcome. In DATA 1800 and DATA 2800, student projects were collected and related to the program learning outcome.

DATA 1800 and DATA 2800 were chosen because they are the most advanced courses that our majors have taken that are specific to data science. In AY 2021-2022, we will compare the results of the assessment in these introductory courses to the results of the assessment of our first graduating class's work in DATA 4961/4962, which is our capstone course.

STAT 3850 is a crucial course in the curriculum of data science. It is a prerequisite for all further study in statistics as well as machine learning.

STAT 4840 is an example of an advanced course that has STAT 3850 as a prerequisite.

The learning outcomes that were chosen were:
DATA 1800: Graduates will be able to use technical skills in predictive modeling.
DATA 2800: Graduates will be able to use technical skills in predictive modeling.
STAT 3850: CLO: perform and interpret single and multiple regression using R, related to PLO Graduates will apply statistics to analyze data sets.
STAT 3850: CLO: recognize the type of random variable that an experiment describes, related to PLO Graduates will apply statistics to analyze data sets.
STAT 4840: CLO: visualize time series, including trend, seasonal and random components, related to PLO Graduates will use programming and other computer science skills to analyze and interact with data.

3. Assessment Methods: Evaluation Process

What process was used to evaluate the artifacts of student learning, and by whom? Please identify the tools(s) (e.g., a rubric) used in the process and **include them in/with this report document** (please do not just refer to the assessment plan).

The rubric used to assess the artifact is given below:

0: Student shows little or no understanding of the concept(s)
1: Student shows a limited understanding of the concept(s)
2: Student shows competence, but not complete mastery of the concept(s)
3: Student shows mastery of the relevant concept(s)

The artifact was assessed by the course instructor.

4. Data/Results

What were the results of the assessment of the learning outcome(s)? Please be specific. Does achievement differ by teaching modality (e.g., online vs. face-to-face) or on-ground location (e.g., STL campus, Madrid campus, other off-campus site)?

See attached

5. Findings: Interpretations & Conclusions

What have you learned from these results? What does the data tell you?

See attached

6. Closing the Loop: Dissemination and Use of Current Assessment Findings

A. When and how did your program faculty share and discuss these results and findings from this cycle of assessment?

Data science faculty met to discuss assessment in Spring 2021. Many results were not yet available, but we discussed the results that we received to that point.

B. How specifically have you decided to use these findings to improve teaching and learning in your program? For example, perhaps you've initiated one or more of the following:

Changes to the Curriculum or Pedagogies	• Course content • Teaching techniques • Improvements in technology • Prerequisites	• Course sequence • New courses • Deletion of courses • Changes in frequency or scheduling of course offerings
Changes to the Assessment Plan	• Student learning outcomes • Artifacts of student learning • Evaluation process	• Evaluation tools (e.g., rubrics) • Data collection methods • Frequency of data collection

Please describe the actions you are taking as a result of these findings.

We are considering changes to the assessment plan. The main issue we face when interpreting the data is that the rubric may not be interpreted the same way by all of the faculty members. We are contemplating providing more instructions in the rubric, which will give guidance to the faculty members when assessing the artifacts of student learning.

If no changes are being made, please explain why.

We need more data in order to make changes. We have not yet taught the entire curriculum to our majors, and we are graduating our first class in Spring 2022. We would like to graduate at least 2 classes before contemplating significant changes.

7. Closing the Loop: Review of Previous Assessment Findings and Changes

A. What is at least one change your program has implemented in recent years as a result of assessment data?

Program started in AY 2019-2020, and our first assessment report was generated in December, 2020. We have made no changes based on previous assessments.

B. How has this change/have these changes been assessed?

C. What were the findings of the assessment?

D. How do you plan to (continue to) use this information moving forward?

IMPORTANT: Please submit any assessment tools (e.g., rubrics) with this report as separate attachments or copied and pasted into this Word document. Please do not just refer to the assessment plan; the report should serve as a stand-alone document.

Supplement to Program Level Assessment Report

Data was collected from the courses DATA 1800, DATA 2800, STAT 3850 (two sections), and STAT 4840. Instructors assigned scores of 0, 1, 2, or 3 based on performance on a course project (in DATA courses) and on a problem on the final exam (in STAT courses). Data is presented in Table 1.

Table 1: Summary Data for Assessment

	Assesment				Summary	
	0	1	2	3	mean	sd
DATA 1800	0	4	6	0	1.60	0.52
DATA 2800	0	1	9	0	1.90	0.32
STAT 3850	0	1	3	41	2.89	0.38
STAT 3850	3	4	8	10	2.00	1.04
STAT 4840	0	0	2	8	2.80	0.42

We see that, as expected, the DATA 2800 students were better overall than the DATA 1800 students at the same program learning objective. Analysis of this data is complicated by the fact that DATA 2800 is expected to be taken by Sophomores, but is being taken by many Juniors who switched into the data science major. The same learning objective will be assessed this year in DATA 4961/4962[1], which will allow us to see wheter the students continue the improvement shown from 1800 to 2800.

For STAT 3850, a Wilcoxon rank sum test with continuity correction shows that there is a significant ($p < 10^{-5}$) difference between the outcomes associated with learning objective "Recognize the type of random variable that an experiment describes" and learning objective "Perform and interpret single and multiple regression using R." Vargha and Delaney's A is 76%, which indicates that a randomly selected student would do better on the regression outcome than the random variable outcome 76 percent of the time. This is surprising because regression is a considerably more challenging topic, which is revisited in a course devoted to the material. One significant issue is that the learning outcome is confounded with instructor, so more data will be needed to determine whether action should be taken.

Meanwhile, the students in STAT 4840, which is the most advanced course assessed in this cycle, did very well, showing near complete mastery of the learning objective in this context.

Finally, we provide a visualization of the data across the courses. We note that the outcomes seem to be improving with the level of the course. We are pleased that our students seem to be becoming more proficient in the learning outcomes as they progress through the program.

[1]DATA 4961/4962 is being offered for the first time this academic year.


Assessment Outcomes by Course
Outcomes improve with level of course
Count
Assessment Score
DATA 1800
DATA 2800
STAT 3850
STAT 3850
STAT 4840
0
1
2
3