Assurance of Student Learning Reflection **2024-2025**	
Ogden College of Science Engineering	*SEAS*
Scientific Data Analytics Certificate (0496)	
Zhonghang Xia	
Is this an online program? ☐ Yes ☒ No	Please make sure the Program Learning Outcomes listed match those in CourseLeaf. Indicate verification here ☒ Yes, they match! (If they don't match, explain on this page under **Evaluation)**

Instructions: For the 2024-25 assessment, we are asking you to reflect on the last three-year cycle rather than collect data. It's important to take time to look over the results from the last assessment cycle and really focus on a data-informed direction going forward. In collaboration with your assessment team and program faculty, review each submitted template from 2021-2024 and consider the following for each Program Learning Outcome, add your narrative to the template, and submit the draft to your ASL Rep by May 15, 2025.

Program Student Learning Outcome 1	
Program Student Learning Outcome	**Write computer programs to utilize and analyze large datasets.**
Evaluation	Using the last three assessment cycles, is this program learning outcome still relevant, or should it be changed? If it has recently changed, please explain. Other things to examine: Is the outcome measurable? Is it double or triple barreled? Does it include measurable verbs following Bloom's Taxonomy? Do you have the appropriate numbers of SLOs to measure regularly? Please consider choosing the most important. Yes. The outcome remains relevant as it addresses foundational data analysis skills in Python, which are critical in today's data-driven job market. No change is needed. The outcome has remained consistent in scope and focus across all three assessment years. Rubrics (O1, O1-1, O1-2) offer clear performance benchmarks. The outcome is clearly measurable. The tasks (data loading, cleaning, analysis, output) are observable, gradable, and aligned with the outcome. Students' abilities are assessed through their code functionality, data cleaning accuracy, analytical processing, and file output—all observable and gradable via rubrics. To some degree, it could be considered double-barreled, since it assesses multiple skills in one outcome (e.g., reading files, cleaning data, analyzing and presenting results). Verbs like "write," "check," "compute," and "select align well with the Application and Analysis levels of Bloom's Taxonomy. There are four SLOs. The outcome reflects key skills in upper-level coursework and capstone-type projects. Outcome consistently meets or exceeds the 80% target benchmark.
Measurement Instruments	Are the measurement instruments actually measuring the outcome? If you change the SLO, is this still the best instrument to use? Is this a direct or indirect measure? Is your artifact appropriate? If not, what other options are there? Will the rise in the use of AI affect the assignment and measurement? If there are rubrics, do they need to be altered to better fit the learning outcome? Does the rubric (if using) work or does it need to be adjusted? The instruments require students to: 1) Write Python programs 2) Clean and manipulate large datasets 3) Apply data analytics and modeling 4) Evaluate and present results. These tasks directly align with the outcome's focus on both programming and data analysis.

	If the SLO is split (e.g., into "Write programs" and "Analyze datasets"), the current assignments may still work but may need re-alignment. It is a directly measure. Students submit Python programs that are evaluated using detailed rubrics. Their work is assessed based on demonstrated skills. The artifact is appropriate. The assignments are authentic, real-world tasks (e.g., cleaning and analyzing movie ratings or air pollution data) and reflect both industry practices and course learning goals. The rise in the use of AI likely affect the assignment and measurement. AI tools (e.g., ChatGPT, GitHub Copilot) could generate code for students, reducing the reliability of code-only assessments. Rubrics O1, O1-1, and O1-2 are appropriate but could be updated to reflect use of AI-assistance guidelines
Criteria & Targets	Does Criteria for Success (level of performance students will have achieved for your program to have been successful--ex., students will have earned 4/5 for documentation and citation on capstone essays) need to be changed? What about targets? If you have successfully made your targets consistently, consider a more challenging target. Criteria for Success are not necessarily need to be changed.
Results & Conclusion	**Results**: Are the results what was expected or not? What stood out in the assessment cycle over the past three years? Explain Yes. All years met the target. The most recent year saw 100% success. Stable achievement across years, despite slight cohort size variation. Consistent use of real datasets with practical relevance. The evaluation results show that teaching methods and content delivery have been effective. **Conclusions**: What worked? What didn't? Why do you think this? For example, maybe the content in one or more courses was modified; changed course sequence (detail modifications); changed admission criteria (detail modifications); changed instructional methodology (detail modifications); changed student advisement process (detail modifications); program suspended; changed textbooks; facility changed (e.g. classroom modifications); introduced new technology (e.g. smart classrooms, computer facilities, etc.); faculty hired to fill a particular content need; faculty instructional training; development of a more refined assessment tool. The wored items include: 1) Python-based assessments using real datasets 2) Rubric-guided evaluations (O1, O1-1, O1-2) 3) Alignment with program and industry needs. 4) Focused course (CS 555) as a dedicated platform for SLO 1. Rubrics or assignments not evolving in light of AI trends.
****IMPORTANT - Plans for Next Assessment Cycle**:	As we work hard to improve our assessment practices and make them more meaningful and effective, it's important each program craft a three-year plan for the following assessment cycle (2025-26, 2026-27, 2027-28) – this process assists in "closing the loop." For example, you may decide to: • collect a more appropriate artifact • create new program outcomes • adjust targets because they are consistently exceeded or not met • need to reconstruct your curriculum map • sequencing of classes might need to be adjusted, or additional class(es) provided Whatever your plan is, provide a narrative, in future tense, that indicates how you will approach future assessments. You will be expected to implement any needed changes before the next assessment cycle. As we continue to refine and strengthen our assessment practices, the following plan outlines targeted actions and improvements for the upcoming three-year cycle. 2025-26 Year: We will revise the current rubrics (O1, O1-1, O1-2) to better account for the rise of generative AI tools (e.g., ChatGPT, Copilot). These revisions will include criteria that assess students' original understanding, such as: Code annotation and justification, In-class coding

	assessments. A new assessment artifact will be piloted: a Jupyter Notebook submission combining code, analysis, and markdown-based interpretation. This format better evaluates students' critical thinking and understanding, not just code output**.** **2026-2027 Year:** We will evaluate whether to split the current outcome into two distinct outcomes: 1. Writing programs to process large datasets 2. Analyzing and interpreting results using statistical or machine learning models **2027-2028 Year:** We will finalize and implement new SLOs and adjust our program success targets.

Program Student Learning Outcome 2	
Program Student Learning Outcome	**Understand the statistical approaches taken when dealing with large sample sizes.**
Evaluation	Using the last three assessment cycles, is this program learning outcome still relevant, or should it be changed? Other things to examine: Is the outcome measurable? Is it double or triple barreled? Does it include measurable verbs following Bloom's Taxonomy? Yes. It is still relevant. Understanding statistical methods for large sample sizes is **fundamental** in data science, machine learning, and research analysis. This remains highly relevant as data volumes continue to grow. The outcome is measurable. Students are required to apply statistical techniques, explain model performance using metrics like accuracy, precision, recall, and interpret statistical outcomes, such as p-values and confidence intervals. These are all **observable, testable skills** that reflect whether students truly grasp the underlying statistical concepts. It is not double barreled. It focuses on a single concept: understanding statistical methods. However, if assignments begin covering model-based inference (e.g., logistic regression and hypothesis testing), splitting into two SLOs might be appropriate later. The current verb "understand" in SLO2 could be revised to "interpret," which is a higher-level, measurable verb recommended by Bloom's Taxonomy.
Measurement Instruments	Are the measurement instruments actually measuring the outcome? If you change the SLO, is this still the best instrument to use? Is this a direct or indirect measure? Is your artifact appropriate? If not, what other options are there? Will the rise in the use of AI affect the assignment and measurement? If there are rubrics, do they need to be altered to better fit the learning outcome? Does the rubric (if using) work or does it need to be adjusted? The instruments measured the outcome. Each year, students use statistical models or inferential methods such as logistic regression for classification, hypothesis testing and confidence intervals to compare means. These assessments clearly measure understanding and application of statistical approaches to large datasets. If the verb is changed to apply or interpret, the same instruments remain valid, especially when rubrics emphasize: correct method selection, interpretation of results, and statistical accuracy. These measures are direct. Students' work is evaluated via coding output and rubric-based analysis of statistical interpretation, not self-reported perceptions. The artifacts are approparite. Python-based statistical analysis align with the program's technical focus. They simulate realistic data science tasks and are based on authentic, multi-variable datasets.

	The rise in the use of AI potentially affect the assignment and measurement. Students might use AI tools (e.g., ChatGPT) to auto-generate hypothesis tests or code explanations
Criteria & Targets	Does Criteria for Success (level of performance students will have achieved for your program to have been successful (ex., students will have earned 4/5 for documentation and citation on capstone essays) need to be changed? What about targets? No change is needed.
Results & Conclusion	**Results**: Are the results what was expected or not? What stood out in the assessment cycle over the past three years? Explain The results are what was expected. What stood out over the past three assessment cycles was the consistent success of students, with a gradual improvement in performance each year. There was also a notable shift in assessment focus—from model-based evaluations, such as logistic regression, to inferential statistical methods, including hypothesis testing and confidence intervals. Additionally, all measurement instruments were grounded in real-world datasets, which significantly reinforced students' data literacy and ability to interpret statistical results in practical contexts. **Conclusions**: What worked? What didn't? Why do you think this? For example, maybe the content in one or more courses was modified; changed course sequence (detail modifications); changed admission criteria (detail modifications); changed instructional methodology (detail modifications); changed student advisement process (detail modifications); program suspended; changed textbooks; facility changed (e.g. classroom modifications); introduced new technology (e.g. smart classrooms, computer facilities, etc.); faculty hired to fill a particular content need; faculty instructional training; development of a more refined assessment tool. What worked well included the use of Python-based assessments with large datasets, which provided students with practical, hands-on experience. The consistent application of Rubric O2 contributed to reliable evaluation across cohorts. Additionally, the inclusion of multiple inferential techniques and model evaluation metrics enhanced the depth and rigor of statistical learning. However, there are areas that need review. The learning outcome uses the verb "understand," which is not directly measurable and should be revised to a more observable action. Furthermore, Rubric O2 may need to be updated to include guidance on the appropriate use of AI tools and to add criteria specifically targeting students' interpretation and reasoning skills.
****IMPORTANT - Plans for Next Assessment Cycle**:	As we work hard to improve our assessment practices and make them more meaningful and effective, it's important each program craft a three-year plan for the following assessment cycle (2025-26, 2026-27, 2027-28) – this process assists in "closing the loop." For example, you may decide to: • collect a more appropriate artifact • create new program outcomes • adjust targets because they are consistently exceeded or not met • need to reconstruct your curriculum map • sequencing of classes might need to be adjusted, or additional class(es) provided Whatever your plan is, provide a narrative, in future tense, that indicates how you will approach future assessments. You will be expected to implement any needed changes before the next assessment cycle. The following plan outlines targeted actions and improvements for the upcoming three-year cycle. 2025-26 Year:

	We focus on learning outcome revision and rubric alignment. Revise SLO 2 from “Understand…” to “Apply and interpret statistical approaches when analyzing large sample sizes.” Update rubric O2 to emphasize Avoidance of AI misuse (add criteria for originality and reasoning) 2026-27 Year: We plan to enhance existing assignments with a small but powerful requirement: an interpretation section. Students are required to submit a short written explanation alongside their statistical analysis (logistic regression or hypothesis tests). They will explain what statistical method they used and why it was chosen. 2027-28 Year: This year focuses on a light internal review of performance trends. We will confirm whether the revised SLO and updated rubric remain aligned and useful.

Program Student Learning Outcome 3	
Program Student Learning Outcome	**Understand the statistical approaches taken when dealing with multiple variables.**
Evaluation	Using the last three assessment cycles, is this program learning outcome still relevant, or should it be changed? Other things to examine: Is the outcome measurable? Is it double or triple barreled? Does it include measurable verbs following Bloom’s Taxonomy? It is still relevant. Understanding statistical methods for analyzing multiple variables is fundamental in data science, particularly for machine learning, multivariate analysis, and real-world decision-making. It is measurable. Like SLO 2, the verb *“understand”* is abstract. However, the use of coding assignments and evaluation rubrics transforms this into measurable actions (e.g., model construction, evaluation, and interpretation). It is not double barreled. It focuses on a single cognitive skill: analyzing relationships between variables using statistical techniques.
Measurement Instruments	Are the measurement instruments actually measuring the outcome? If you change the SLO, is this still the best instrument to use? Is this a direct or indirect measure? Is your artifact appropriate? If not, what other options are there? Will the rise in the use of AI affect the assignment and measurement? If there are rubrics, do they need to be altered to better fit the learning outcome? Does the rubric (if using) work or does it need to be adjusted? The measurement instruments are measuring the outcome. Students build and evaluate models (e.g., tree-based regressors, scikit-learn pipelines) that analyze the influence of multiple predictors (e.g., PM10, CO, SO_2, NO_x) on a target variable (PM2.5). This directly supports the outcome. If the outcome is rewritten to use “apply” or “evaluate,” the current machine learning model assignment still fits well—especially when paired with Rubric O3. They are direct measures. Students’ submissions and model evaluation reports are tangible evidence of learning. Yes, the artifiact is appropriate.The dataset and assignment are realistic and require application of multivariate analysis using programming and data science tools. AI tools (e.g., ChatGPT, Copilot) can generate code and explanations. We can ask students to justify model selection and include interpretation or reflection components.
Criteria & Targets	Does Criteria for Success (level of performance students will have achieved for your program to have been successful (ex., students will have earned 4/5 for documentation and citation on capstone essays) need to be changed? What about targets? No change is needed.

Results & Conclusion	**Results**: Are the results what was expected or not? What stood out in the assessment cycle over the past three years? Explain What stood out over the past three assessment cycles was that student performance consistently exceeded expectations, particularly in the most recent two cycles. There was also a clear progression in the complexity of modeling tasks, moving from tree-based models to more flexible and advanced model options. Throughout the assessments, students worked with real-world datasets containing multiple features, which helped reinforce practical data analysis skills. Most notably, students consistently demonstrated the ability to manage and interpret multiple predictors effectively in their modeling tasks. **Conclusions**: What worked? What didn't? Why do you think this? For example, maybe the content in one or more courses was modified; changed course sequence (detail modifications); changed admission criteria (detail modifications); changed instructional methodology (detail modifications); changed student advisement process (detail modifications); program suspended; changed textbooks; facility changed (e.g. classroom modifications); introduced new technology (e.g. smart classrooms, computer facilities, etc.); faculty hired to fill a particular content need; faculty instructional training; development of a more refined assessment tool. What worked well was the use of Python-based modeling assignments involving large datasets, which provided students with practical experience in applying statistical methods. The consistent application of Rubric O3 helped ensure uniform evaluation of student work across cohorts. Additionally, the use of scikit-learn tools supported realistic model development and evaluation, aligning closely with industry practices. However, one area that needs improvement is the use of the verb "understand" in the learning outcome, as it is not directly observable or easily measurable. A revision to a more action-oriented verb is recommended.
****IMPORTANT - Plans for Next Assessment Cycle**:	As we work hard to improve our assessment practices and make them more meaningful and effective, it's important each program craft a three-year plan for the following assessment cycle (2025-26, 2026-27, 2027-28) – this process assists in "closing the loop." For example, you may decide to: • collect a more appropriate artifact • create new program outcomes • adjust targets because they are consistently exceeded or not met • need to reconstruct your curriculum map • sequencing of classes might need to be adjusted, or additional class(es) provided Whatever your plan is, provide a narrative, in future tense, that indicates how you will approach future assessments. You will be expected to implement any needed changes before the next assessment cycle. 2025-26 Year: In the first year, we will revise Student Learning Outcome 3 to use a more measurable verb. This change will make the outcome more actionable and aligned with Bloom's Taxonomy. Alongside this revision, Rubric O3 will be updated to include clearer criteria for evaluating students' ability to select appropriate models, interpret the influence of multiple predictors, and address concerns such as overfitting and generalizability. 2026-27 Year: In the second year, we will enhance existing assessments by adding a brief interpretation component. Students will be required to include a short written explanation with their model output. This explanation will describe their rationale for model selection, interpret the relationships among variables, and discuss key findings. Feature importance plots or residual analysis may also be encouraged to support interpretation. 2027-28 Year: In the third year, we will conduct a review of rubric scores and assessment results from the previous two cycles to identify trends in student performance. This evaluation will help determine whether students are consistently able to explain relationships between multiple variables and apply appropriate statistical approaches.

Program Student Learning Outcome 4	
Program Student Learning Outcome	**Combine domain expertise with programming and statistical skills to analyze large domain-specific datasets.**
Evaluation	Using the last three assessment cycles, is this program learning outcome still relevant, or should it be changed? Other things to examine: Is the outcome measurable? Is it double or triple barreled? Does it include measurable verbs following Bloom's Taxonomy? SLO 4 remains highly relevant in today's data-driven environment, where real-world problems often require integrating domain knowledge (e.g., air quality, finance, health) with machine learning and statistical tools. It supports interdisciplinary data science applications—an essential skill for modern graduates. The outcome is clearly measurable through direct assessments of students' ability to 1)apply machine learning algorithms 2)analyze results with statistical metrics 3)interpret domain-specific patterns (e.g., air pollution trends). It is double barreled. It combines domain expertises, programming skills, and statistical analysis. "Combine" can be changed to "apply".
Measurement Instruments	Are the measurement instruments actually measuring the outcome? If you change the SLO, is this still the best instrument to use? Is this a direct or indirect measure? Is your artifact appropriate? If not, what other options are there? Will the rise in the use of AI affect the assignment and measurement? If there are rubrics, do they need to be altered to better fit the learning outcome? Does the rubric (if using) work or does it need to be adjusted? Yes. Across all assessment cycles, students consistently built multiple machine learning models to analyze large datasets. They evaluated the performance of these models using metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Receiver Operating Characteristic (ROC) curves. To interpret the contribution of individual features to model predictions, students utilized tools such as SHAP values and partial dependence plots. All assignments were based on realistic, domain-specific datasets, such as those related to air pollution. These tasks collectively provided a strong measure of students' ability to integrate domain knowledge with programming and statistical analysis. It is a direct measure. Students' projects and reports provide observable evidence of learning. It is appropriate. The large dataset, model comparison, and interpretation tasks are all authentic, appropriate, and aligned with the outcome. Students might use AI (e.g., Copilot) to auto-generate code. Rubric O4 works well but could be improved by adding a section to address responsible AI/tool use.
Criteria & Targets	Does Criteria for Success (level of performance students will have achieved for your program to have been successful (ex., students will have earned 4/5 for documentation and citation on capstone essays) need to be changed? What about targets? No change is needed.
Results & Conclusion	**Results**: Are the results what was expected or not? What stood out in the assessment cycle over the past three years? Explain

	Yes. What stood out during the past three assessment cycles was the consistent performance of students, who regularly met or exceeded expectations. There was a clear progression in assignment complexity—from comparing just two models to evaluating multiple machine learning algorithms using various performance metrics. The datasets used throughout remained realistic and domain-specific, such as air quality data with multiple input variables, which enhanced the relevance of the analysis. Additionally, students demonstrated strong skills in applying both statistical evaluation metrics and visual interpretation tools, such as SHAP values and partial dependence plots, to explain their model results. **Conclusions**: What worked? What didn't? Why do you think this? For example, maybe the content in one or more courses was modified; changed course sequence (detail modifications); changed admission criteria (detail modifications); changed instructional methodology (detail modifications); changed student advisement process (detail modifications); program suspended; changed textbooks; facility changed (e.g. classroom modifications); introduced new technology (e.g. smart classrooms, computer facilities, etc.); faculty hired to fill a particular content need; faculty instructional training; development of a more refined assessment tool. What worked well across the assessment cycles was the effective use of Python to analyze large, real-world datasets, which provided students with valuable hands-on experience. The consistent application of Rubric O4 ensured reliable and standardized evaluation of student performance. Assessments were based on a well-balanced combination of model performance metrics and interpretative analysis, allowing students to demonstrate both technical and analytical skills. Furthermore, there was a clear progression in assignment complexity, evolving from comparisons of two models in earlier cycles to more advanced tasks involving multiple models and robust performance evaluations. The SLO verb "combine" could be clearer and more actionable.
****IMPORTANT - Plans for Next Assessment Cycle**:	As we work hard to improve our assessment practices and make them more meaningful and effective, it's important each program craft a three-year plan for the following assessment cycle (2025-26, 2026-27, 2027-28) – this process assists in "closing the loop." For example, you may decide to: • collect a more appropriate artifact • create new program outcomes • adjust targets because they are consistently exceeded or not met • need to reconstruct your curriculum map • sequencing of classes might need to be adjusted, or additional class(es) provided Whatever your plan is, provide a narrative, in future tense, that indicates how you will approach future assessments. You will be expected to implement any needed changes before the next assessment cycle. 2025-26 Year: In the first year, the program will revise the wording of Student Learning Outcome 4 to enhance clarity and measurability. This change will align the outcome more closely with observable student actions and better reflect expectations in data science. Additionally, Rubric O4 will be updated to include specific criteria for evaluating how well students integrate domain knowledge with model selection, statistical interpretation, and programming accuracy. The revised rubric will also distinguish between intermediate and advanced levels of performance more clearly. 2026-27 Year: In the second year, the program will strengthen the interpretative aspect of the assessment. Students will be required to include a short written section in their project reports explaining how they selected their machine learning models based on domain-specific factors and how key variables influenced the predictions. This explanation will also reference model outputs such as feature importance (e.g., SHAP values) or residual plots, connecting the technical results back to the real-world context of the dataset. This adjustment will support deeper critical thinking and ensure that students are not only generating accurate models but also interpreting them meaningfully. 2027-28 Year:

	In the third year, the program will conduct a review of student performance data using Rubric O4 from the previous two assessment cycles. This review will identify trends in students' abilities to combine domain knowledge with statistical and programming skills. Based on these insights, faculty will determine whether instructional adjustments are needed—such as providing more guidance on model justification or interpretation techniques.

To add more outcomes, if needed, select the table above and copy & paste below.