

**PROGRAM ASSESSMENT REPORT**

**MASTER OF SCIENCE IN**

**DATA SCIENCE AND ENGINEERING**

The Master of Science in Data Science and Engineering (MS-DSE) degree entails a minimum of 30 semester graduate credit hours. There are two formal options in the MS-DSE program: the *Data Engineering Option* administrated by the Department of Computer Science and Software Engineering; and the *Data Science Option* managed by the Department of Mathematics and Statistics. The course list for the data engineering option:

|  |  |  |
| --- | --- | --- |
| [COMP 6120](https://nextbulletin.auburn.edu/search/?P=COMP%206120) | Database Systems I | 3 |
| [COMP 6130](https://nextbulletin.auburn.edu/search/?P=COMP%206130) | Data Mining | 3 |
| [COMP 6630](https://nextbulletin.auburn.edu/search/?P=COMP%206630) | Machine Learning | 3 |
| [STAT 6000](https://nextbulletin.auburn.edu/search/?P=STAT%206000) | Intermediate Statistical Methods for Data Science | 3 |
| [STAT 6600](https://nextbulletin.auburn.edu/search/?P=STAT%206600) | Probability and Statistics for Data Science | 3 |
| [STAT 6650](https://nextbulletin.auburn.edu/search/?P=STAT%206650) | Statistical Learning | 3 |
| Select 6 Credits in [COMP 6000](https://nextbulletin.auburn.edu/search/?P=COMP%206000)-8999 | | 6 |
| Select 3 Credits in 6000-8999 | | 3 |
| [COMP 7980](https://nextbulletin.auburn.edu/search/?P=COMP%207980) | Capstone Engineering Project | 3 |
| **Total** | | **30** |

The course list for the data science option:

|  |  |  |
| --- | --- | --- |
| [STAT 6000](https://nextbulletin.auburn.edu/search/?P=STAT%206000) | Intermediate Statistical Methods for Data Science | 3 |
| [STAT 6600](https://nextbulletin.auburn.edu/search/?P=STAT%206600) | Probability and Statistics for Data Science | 3 |
| [STAT 6650](https://nextbulletin.auburn.edu/search/?P=STAT%206650) | Statistical Learning | 3 |
| [COMP 6120](https://nextbulletin.auburn.edu/search/?P=COMP%206120) | Database Systems I | 3 |
| [COMP 6130](https://nextbulletin.auburn.edu/search/?P=COMP%206130) | Data Mining | 3 |
| [COMP 6630](https://nextbulletin.auburn.edu/search/?P=COMP%206630) | Machine Learning | 3 |
| Select 6 Credits in [STAT 6000](https://nextbulletin.auburn.edu/search/?P=STAT%206000)-8999 | | 6 |
| Select 3 Credits in 6000-8999 | | 3 |
| [STAT 7940](https://nextbulletin.auburn.edu/search/?P=STAT%207940) | Capstone Project | 3 |
| **Total** | | **30** |

Depending on the option, graduate students are required to take the following courses. Two formal options share the same list of six foundational courses. Among the three elective courses related to data science and engineering, Data Engineering-option students must choose at least two COMP6/7000-level graduate courses, whereas Data Science-option students are required to take at least two STAT6/7000-level graduate courses.

The two options offer a list of elective courses that are highly relevant to data science and engineering. The course selections are at the discretion of the students and the directors of graduate programs (a.k.a., GPOs).

Students who have taken the 5000‐level version of any 6000‐level course cannot take the course in its 6000‐level format. Those students must substitute other graduate courses as directed by the GPOs. All the courses including COMP7980 and STAT7940 in the degree must be taken for a grade.

Student Learning Outcomes

This graduate program prepares students to pursue careers in data science and engineering, where valuable insights are derived from massive amounts of raw data. Our high‐quality curriculum offers an excellent balance between theory and application, equipping students with foundational skills and state‐of‐the‐art technologies related to the next generation of big data applications. Students will be able to complete this on‐campus graduate program in one to two years. This program blends graduate‐level courses in core topics like data mining, machine learning, and statistical learning. The program also offers a wide variety of electives in addition to a required capstone experience, in which students apply their knowledge and skills to a real‐world application scenario.

1. Specificity of Outcomes

The MS-DSE program encompasses four student learning outcomes (SLO 1-4).

* **SLO 1:** Students will incorporate the theoretical math and statistical principles to design and implement solutions to solve data analytics problems.
* **SLO 2:** Students will utilize practical computing technologies to design and implement data-driven solutions.
* **SLO 3:** Students will apply data science and engineering methods to at least one domain of real-world applications.
* **SLO 4:** Students will deliver written and oral presentations to non-technical and technical audiences during the design and implementation phases of data science and engineering projects.

2. Comprehensive Outcomes

The four student learning outcomes (SLOs) listed in Section 1 represent the culmination of all coursework related to the MS-DSE program. These outcomes represent skills and capabilities that should be fully developed over the course of each MS-DSE student’s education.

Although the program offers a broad education allowing MS-DSE students to equip with data mining, machine learning, and statistical learning skills, these SLOs represent the skills and capacities required in service as data science and engineering professionals.

These student learning outcomes are in alignment with disciplinary standards set by the American Statistical Association (ASA), the Association of Computing Machinery (ACM), and those professionals represented by ASA and ACM.

The data science and engineering committee (Drs. Zhou, Karmaker, Ku, and Qin) at the CSSE department originated an earlier version of the SLOs. After the SLOs reviewed by the CSSE graduate programs committee, the SLOs are further revised by the joint MS-DSE committee at CSSE and DMS. In June 2020, the SLOs were approved by the CSSE and DMS faculty. The program’s four SLOs will be re-evaluated, reviewed, and ratified in the year 2025.

3. Communicating Student Learning Outcomes

3.1 Faculty

The graduate programs committee was directly involved in the development of program vision and student learning outcomes. A copy of this assessment report will be disseminated to newly hired faculty and adjunct faculty members during the orientation week or beginning of their first academic semester. We have made all the assessment reports available on the department’s shared drive to be accessed by all the CSSE and DMS faculty and staff members.

3.2 Students

We have provided newly admitted MS-DSE students with a copy of program vision and student learning outcomes in their orientation meeting organized by the Graduate Program Officer. The orientation meetings are held in the beginning of the first semester of the program. MS-DSE Students are introduced about the program assessment methods used by the graduate program committee. The student learning outcomes are posted on the departmental website to disseminate the SLOs information to current and prospective MS-DSE students.

Curriculum Map

4. Curriculum Map for the MS-DSE program

The following table is a curriculum map that visually depicts the alignment between the MS-CSSE student learning outcomes (SLOs) and the courses offered from the Fall’19 to Spring’20 semesters. In each cell of the curriculum map, a letter indicates how a course relates to a student learning outcome. Letters “I”, “R”, and “E” designate courses in the program “introduce”, “reinforce”, or “emphasize” the corresponding learning outcomes.

**Contribution Weight Notation**

* Introduce: 0.33 I
* Reinforce: 0.66 R
* Emphasize: 1.0 E

The curriculum map was reviewed and approved by the CSSE faculty in June 2020. The graduate programs committee intends to keep the curriculum maps unchanged in the next five years unless one or more learning outcomes are updated before June 2025. We will revisit and update this curriculum map in the spring 2025 semester.

The Curriculum Map of the MS-DSE Program

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Course Outcomes** | **SLO 1** | **SLO 2** | **SLO 3** | **SLO 4** |
| STAT 6000/6 Intermediate Statistical Methods for Data Sci | E |  | I |  |
| STAT 6600/6 Probability and Statistics for Data Science | E |  | I |  |
| STAT 6650/6 Statistical Learning | E |  | I |  |
| STAT 7940/6 Capstone Project | R | R | E | E |
| COMP 6120/6 Database Systems I |  | E | I | R |
| COMP 6130/6 Data Mining | R | E | I | R |
| COMP 6630/6 Machine Learning | R | E | I | I |
| COMP 7980/6 Capstone Engineering Projects | R | R | E | E |
| STAT 7030 Categorical Data Analysis | E | R | R | R |
| COMP 6000/6 Web Application Development |  | R | R |  |
| COMP 6320/6 Design and Analysis of Computer Networks |  | R | E |  |
| COMP 6340/6 Network Quality Assurance and Simulation |  | R | E |  |
| COMP 6350/6 Digital Forensics |  | R | R |  |
| COMP 6360/6 Wireless and Mobile Networks |  | R | E |  |
| COMP 6370/6 Computer and Network Security |  |  | E |  |
| COMP 6400/6 Foundation of Computer Graphics |  |  | E |  |
| COMP 6520/6 Network and Operating Sys Admin |  |  | E |  |
| COMP 6600/6 Artificial Intelligence |  | E |  |  |
| COMP 6620/6 User Interface Design and Evaluation |  |  |  | I |
| COMP 6700/6 Software Process |  | R | R |  |
| COMP 6710/6 Software Quality Assurance |  | R | E |  |
| COMP 6970/6 Special Topics: Comp Int. & Adversarial ML |  | E | R | R |
| COMP 6970/6 Special Topics: Intro to Evolutionary Comp |  | E | R | R |
| COMP 6970/6 Special Topics: Game Design for SC |  |  |  | I |
| COMP 6970/6 Special Topics: Cybersec Threats & CM |  |  |  | I |
| COMP 6970/6 Special Topics: Computational Biology |  |  | R | I |
| COMP 6970/6 Special Topics: Computer Sec at the Fringes |  |  | R |  |
| COMP 6970/6 Special Topics: Deep Learning |  | E |  | R |
| COMP 6970/6 Special Topics: Game Design and Dev |  |  |  |  |
| COMP 6970/6 Special Topics: Information Retrieval |  | E |  | R |
| COMP 6970/6 Special Topics: Sec in Wireless Networks |  |  | R |  |
| COMP 6970/6 Special Topics: Cloud Computing |  | E |  |  |
| COMP 6970/6 Special Topics: Software Analytics |  |  | R |  |
| COMP 7120/6 Database Systems II |  | E |  | R |
| COMP 7270/6 Advanced Topics in Algorithms |  | R |  |  |
| COMP 7300/6 Advanced Computer Architecture |  |  | R |  |
| COMP 7330/6 Topics in Parallel and Distributed Computing |  | R | R | R |
| COMP 7370/6 Advanced Computer and Network Security |  |  | R |  |
| COMP 7500/6 Advanced Topics in Operating Systems |  | R |  |  |
| COMP 7620/6 Human Computer Interaction |  |  | E |  |
| COMP 7700/6 Software Architecture | R | I |  | I |
| COMP 7720/6 Software Re-Engineering |  | I |  |  |
| COMP 7950 Introduction Graduate Study Computer Science |  |  |  | I |
| COMP 7930/6 Directed Study | R | R | E | E |
| STAT6110 SAS Programming and Applications |  | R | E | R |
| STAT6670 Probability and Stochastic Processes I | R |  | E |  |
| STAT6680 Probability and Stochastic Processes II | R |  | E |  |
| STAT7000 Experimental Statistics I | I | E | E | R |
| STAT7010 Experimental Statistics II | I | E | E | R |
| STAT7020 Regression Analysis | I | E | E | R |
| STAT7030 Categorical Data Analysis | I | E | I | R |
| STAT7040 Biostatistics |  | R | E | E |
| STAT7250 Practical Data Analysis and Computation for the Life Sciences |  | R | I | E |
| STAT7620 Nonparametric Statistics | I | E | E | E |
| STAT7650 Computational Statistics |  | R | E | R |
| STAT7670 Applied Longitudinal Data Analysis | I | E | E | R |
| STAT7700 Generalized Linear Models | I | E | I | E |
| STAT7780 Survival Analysis | I | E | I | R |
| STAT7800 Linear Models | R | I | I | R |
| STAT7810 Modern Stochastic Processes I | R |  | E |  |
| STAT7820 Applied Stochastic Processes I | R |  | E |  |
| STAT7830 Applied Stochastic Processes II | R |  | E |  |
| STAT7840 Applied Multivariate Statistical Analysis | I | R |  | R |
| STAT7850 Theory of Statistical Inference | R | R | E | E |
| STAT7860 Applied Time Series Analysis |  |  |  | I |

We translate “introduce”, “reinforce”, or “emphasize” into weights so that we can gauge aggregate direct measurement data from all the contributing graduate courses. More specifically, we assign weights to represent the contributions of the graduate courses to each of the four SLOs. The weights for “introduce”, “reinforce”, or “emphasize” are 0.33, 0.66, and 1.0, respectively.

The Course Outcome Map of the MS-DSE Program

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Course Outcomes** | **SLO 1** | **SLO 2** | **SLO 3** | **SLO 4** |
| STAT 6000 Intermediate Statistical Methods for Data Sci | 1.00 |  | 0.33 |  |
| STAT 6600 Probability and Statistics for Data Science | 1.00 |  | 0.33 |  |
| STAT 6650 Statistical Learning | 1.00 |  | 0.33 |  |
| STAT 7940 Capstone Project | 1.00 | 0.66 | 0.66 | 0.66 |
| COMP 6120 Database Systems I | 0.33 | 1.00 | 1.00 | 0.66 |
| COMP 6120 Database Systems I (Summer) | 0.33 | 1.00 | 0.66 | 1.00 |
| COMP 6130 Data Mining | 0.66 | 1.00 | 0.33 | 1.00 |
| COMP 6360 Wireless and Mobile Networks |  | 0.66 | 1.00 |  |
| COMP 6630 Machine Learning | 0.33 | 0.33 | 0.33 | 0.33 |
| COMP 6360 Wireless and Mobile Networks | 0.33 | 0.33 | 0.66 | 0.33 |
| COMP 6710 Software Quality Assurance | 0.33 | 0.66 | 0.33 | 1.00 |
| COMP 7980 Capstone Engineering Projects | 0.66 | 0.66 | 1.00 | 1.00 |
| COMP 6320 Design and Analysis of Computer Networks | 0.33 | 0.66 | 1.00 | 0.00 |
| COMP 6970 Special Topics: Cyber Physical Systems Security |  |  | 0.33 | 0.33 |
| COMP 6970 Special Topics: iOS Development |  | 0.33 | 0.33 | 0.33 |
| COMP 6970 Special Topics: Information Retrieval |  | 0.66 | 0.33 | 0.33 |
| COMP 7720 Software Re-Engineering | 0.33 |  |  | 0.33 |
| COMP 7970 Natural Language Processing |  | 1.00 | 0.66 | 0.66 |
| STAT 7000 Experimental Statistics I | 1.00 | 1.00 | 1.00 | 0.66 |
| STAT 7030 Categorical Data Analysis | 0.66 | 0.33 | 1.00 | 0.33 |
| Weights | 6.96 | 6.96 | 8.30 | 6.32 |

Measurement

5. Outcome-Measure Alignment

All the student learning outcomes are assessed using a combination of pre-tests and post-tests, quizzes, midterm and final exams, written assignments, programming assignments, course projects, or presentations. A handful of rubrics have been developed to assess homework and project assignments (e.g., written communication, oral communication, projects).

The outcome-measure alignment is listed in the table below. Notation: QZ - Quiz, PE – Pre-test, ME – Midterm Exams, FE – Final Exams, WA – Written Assignments, PA -Programming Assignments, PT – Presentations, PJ – Projects, QF – Qualifying Exam, PD – Proposal Defense, FD – Final Defense, ID – Indirect, SV – Survey.

The outcome-measure alignment of the MS-DSE Program

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Outcomes** | **QZ** | **PE** | **ME** | **FE** | **WA** | **PA** | **PJ** | **PT** | **QF** | **PD** | **FD** | **ID** | **SV** |
| SLO 1 | X |  | X | X | X | X |  |  | X |  |  |  |  |
| SLO 2 |  | X | X | X |  |  | X |  |  |  |  |  |  |
| SLO 3 |  |  |  |  | X |  | X | X |  | X | X |  |  |
| SLO 4 |  |  |  |  |  |  | X |  |  | X | X | X | X |

6. Direct Measures

Each student’s learning outcome has a list of direct measurements. Direct measures, requiring master’s students to demonstrate their knowledge and skills, will provide tangible, visible and self-explanatory evidence of what our graduate students have and have not learned as a result of graduate courses and the master’s program. In this section, we elaborate on various ways of measuring actual student behaviors.

6.1 Authentic Course-Embedded Measures

This type of direct measure includes exams, tests, quizzes, oral presentations, quizzes, assignments. We believe the direct measures in this category are most effective because the measures are course-embedded, which means the work accomplished by our Ph.D. students are work that counts towards a grade. Our Ph.D. students are likely to take the learning activity more seriously if associated with grades. We directly gauge student knowledge and skills by the virtue of authentic and part of already existing student work. The list of direct measures from the graduate courses offered in the MS-DSE program is summarized below.

* Exams
* Quizzes
* Oral Presentations
* Assignments
* Projects.

The disadvantage of this method is two-fold: it is immensely time-consuming, and it is difficult to collect data to forge an annual assessment report. For this reason, we also rely on the following direct measures of student learning.

6.2 Exit Exams: Capstone Data Science and Engineering Projects

The exit examinations, capstone data science and engineering project defenses will take place at the conclusion of a master’s capstone engineering project. The examination will be administered by the candidate's advisory committee in accordance with the Auburn University Bulletin. Before this examination can be taken, the student must have satisfied all the other course requirements for the master’s degree.

7. Data Collection

The graduate faculty members from the Department of Computer Science and Software Engineering and the Department of Mathematics and Statistics collaboratively collect data for the core required courses and elective courses, usually in a wide variety of forms like quizzes, exams, assignments, course projects, and presentations. Examinations consist of a series of questions administered in class at the end of the semester. Quiz and exam answers are reviewed by the faculty members and reported to the assessment coordinator.

In each graduate course, the faculty member aligned at least one of the direct measures - exams, tests, quizzes, oral presentations, quizzes, and assignments - to at least one of the four student learning outcomes. For each direct measure in a graduate course, student performance ratings are categorized in five camps depending on score ranges:

**Direct Measure Rating Score Range Notation**

* Outstanding Mastery: [90, 100] OM
* Above Satisfactory Mastery: [80, 90) AM
* Satisfactory Mastery: [70, 80) SM
* Partial Mastery: [60, 70) PM
* Minimal or no evidence of Mastery: [0, 60) MM

The table below includes all relevant information with regards to how the direct assessment was conducted and the rationale for choosing specific measurement methods.

The Detailed Information on Direct Assessment Methods

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Learning Outcomes | SLO 1 | SLO 2 | SLO 3 | SLO4 |
| Direct Measure  Description | Exams and Quizzes | Programming Assignment and Projects | Assignment and Projects | Presentations and Written Reports |
| Samples | All students enrolled in STAT 6000, 6600, 6650 | All students enrolled in COMP 6120, 6130, 6630 | All students enrolled in COMP 6320, 6340, 6370, 6600, 6620, 6970 | All students enrolled in COMP 7980 and STAT 7940 |
| When | Final exam week; fall and spring semesters | Last week of fall and spring semesters | Last week of fall and spring semesters | Presentation weeks of fall and spring semesters |
| Where | Final Week and In-Class or Canvas | Canvas | Canvas | Final Week and In-Class |
| How | Faculty or GTA graded | Faculty or GTA graded | Faculty or GTA graded | Faculty graded |
| Desired Results | The class average should be rated as “Outstanding Mastery” | The class average should be rated as “Outstanding Mastery” | The class average should be rated as “Outstanding Mastery” | The class average should be rated as “Outstanding Mastery” |

CSSE and DMS graduate faculty members have taught a total of 9 graduate-level courses to MS-DSE students in the academic year 2020-2021. At the end of the spring 2021 semester, the assessment coordinator (CSSE GPO) designed a course-level data collection form shared on the Google Spreadsheet. After the data-collection requests were sent to all the MS-DSE instructors, the data collection spreadsheet was collaboratively updated in a timely manner.

The assessment data for SLO 1 are gauged from the three STAT core courses, namely, STAT 6000 Intermediate Statistical Methods for Data Sci, STAT 6600 Probability and Statistics for Data Science, and STAT 6650 Statistical Learning. Similarly, the SLO 2 assessment data are contributed by the three required COMP courses, namely, COMP 6120 Database Systems I, COMP 6130 Data Mining, and COMP 6630 Machine Learning. A wide variety of elective courses offer a way of collecting assessment data for SLO 3.

The SLO 4 assessment data related to oral and written communication skills are gathered throughout STAT 7940 Capstone Project and COPM 7980/6 Capstone Engineering Projects. The rubrics are used to assess both written and oral communication skills during project development lifecycles. The faculty members complete the rubric distributed to students to improve their communication performance. All scores for written reports and presentations are collected by the assessment coordinator at the end of each semester.

MS-DSE students with the data science option should complete STAT 7940 capstone project; MS-DSE students with the data engineering option must accomplish COMP 7980 capstone engineering projects. A student’s supervisor completes a performance evaluation form to assess the student’s communication, research, and software development skills. The survey is provided as a Qualtrics survey, which is completed online by STAT and COMP graduate faculty members. To offer feedback to MS-DSE students, the faculty will craft mid-semester reports on the research progress of the students. Final evaluations is intended to be available at the end of each semester collected by the assessment coordinators.

Results

8. Reporting Results

Recall that in each graduate course, the faculty member aligned at least one of the direct measures - exams, tests, quizzes, oral presentations, quizzes, and assignments - to at least one of the four student learning outcomes. Given direct measures in the graduate courses, the instructors rated student performance in the following five categories depending on score ranges.

**Direct Measure Rating Score Range Notation**

* Outstanding Mastery: [90, 100] OM
* Above Satisfactory Mastery: [80, 90) AM
* Satisfactory Mastery: [70, 80) SM
* Partial Mastery: [60, 70) PM
* Minimal or no evidence of Mastery: [0, 60) MM

The score of each course is mapped to the four SLOs according to the curriculum map delineated in Section 4 on page 3. The SLO scores contributed by each graduate course are listed in the table below. It is noteworthy that all the scores recorded in the following table are automatically derived from the curriculum map coupled with the above course-level direct measurement data.

The direct measurement data, the collection method of which are articulated in Sections 5 and 6, are summarized in the following table for the MS-DSE program. The score of each course is mapped to the four SLOs according to the curriculum map delineated in Section 4 on page 3.

The Direct Measurement Data for the MS-DSE Programs

(Part 1/3)

A picture containing text, screenshot, parallel, design

Description automatically generated

The Direct Measurement Data for the MS-DSE Programs

(Part 2/3)

A picture containing text, screenshot, parallel, number

Description automatically generated

The Direct Measurement Data for the MS-DSE Programs

(Part 3/3)

A picture containing text, screenshot, font, number

Description automatically generated

A picture containing text, screenshot, number, parallel

Description automatically generated

We only collected data from the required courses and three electives. The assessment data from directed study, COMP 7980/6 capstone engineering projects, and STAT 7940 capstone project will be collected in the 2020-2021 academic year. Each SLO metric is evaluated in the scale from 100 to 0, where a score in [100, 90] is *Exemplary*, (90, 80] is *Proficient*, (80, 70] is *Acceptable*, and (70, 0] is *Deficient*.

The following table summarizes the SLO results derived from the course-level data. It is noteworthy that all the scores recorded in the following table are automatically derived from the curriculum map coupled with the above course-level grade data.

The SLO Results Derived from the Direct Measurement Data

(Part 1/2)

**A screenshot of a computer

Description automatically generated with low confidence**

The SLO Results Derived from the Direct Measurement Data

(Part 2/2)

**A screenshot of a computer

Description automatically generated with low confidence**

9. Interpreting Results

The aggregate direct measurement data along with the data assessed in the previous cycles are summarized in the data comparison table on page 15 – the next page. The key takeaways from the assessment data are:

* All the four SLOs received ratings of “Exemplary”.
* The highest score goes to SLO 1, which is 97.8.
* With respect to SLOs 2-4, student performance measures in the current cycle has an edge over those of the previous one.
* The student performance measures are leaping across the board in SLOs 1-4.
* The students’ data science skills are superior to their data engineering skills.
* The MS-DSE faculty will continue bolstering student outcomes in terms of SLO2 and SLO4.

Our target scores for all the four student learning outcomes (SLOs) are 90 or above, which represents a rating of “Exemplary”. The final scores for the four SLOs are 97.8, 95.4, 96.9, and 95.0, respectively. Surprisingly, the SLO 2-4 measures gain improvements of 1.0%, 0.3%, and 0.2% over those assessed in the previous cycle. Again, SLO 1 score - the winner among its peers – is higher than that in cycles 2019-2020 and 2020-2021. MS-DSE students are competent in incorporating the theoretical math and statistical principles to design and implement solutions to solve data analytics problems. When it comes to applying data science and engineering methods to at least one domain of real-world applications (SLO 3), our MS-DSE graduate students have been performing very well. The SLO 3 score is surging from 90.9 in the 2020-2021 cycle to 97.0 in the 2020-2021 cycle. The good performance in terms of SLO 1 and SLO 3 implies that our students have gained solid background in the data science discipline. Likewise, The SLO 4 measure jumps from 89.8 to 94.8 – an impressive enhancement of 5.6%.

Direct Measurement Data Comparison between the current cycle (2022-2023) and the previous cycles (2019-2022)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MS-DSE Program – Direct Measures** | **SLO 1** | **SLO 2** | **SLO 3** | **SLO 4** |
| **2022-2023 Final Scores** | **97.8** | **95.4** | **96.9** | **95.0** |
| **Ratings** | **Exemplary** | **Exemplary** | **Exemplary** | **Exemplary** |
| **2021-2022 Final Scores** | 98.1 | 94.5 | 96.6 | 94.8 |
| **Ratings** | Exemplary | Exemplary | Exemplary | Exemplary |
| **2020-2021 Final Scores** | 96.3 | 89.1 | 90.9 | 89.8 |
| **Ratings** | Exemplary | Proficiency | Exemplary | Proficiency |
| **2019-2020 Final Scores** | 93.2 | 92.7 | 86.0 | 95.6 |
| **Ratings** | Exemplary | Exemplary | Exemplary | Exemplary |

Perhaps most importantly, for the first time in the past three years, all the SLO ratings achieve the level of “Exemplary”. We look forward to sharing the encouraging results with the CSSE and DMS graduate faculty and MS-DSE students.

10. Communicating Results

The program outcomes were shared with the CSSE and DMS faculty through emails last year. The program outcomes will be posted on the departmental website within one month after this report is submitted to the Office of Academic Assessment. This report was disseminated to the CSSE faculty members during regular faculty meeting scheduled on June 18, 2023 in the summer 2023 semester. After we incorporated the DSE faculty’s comments on this report, the assessment report is submitted to the Office of Academic Assessment. We also disseminated the report to the DSE committee members after the report’s final submission.

The joint DSE committee members have discussed the assessment findings to offer action items in Section 12 on page 17. The final version of this assessment report will be shared, distributed, and discussed during CSSE and DMS faculty meetings scheduled in the fall 2023 semester. We will share the key takeaways from the assessment report with new MS-DSE during the orientation session scheduled on August 13, 2023.

A picture containing diagram, screenshot, text, circle

Description automatically generated

The above infographic sketches the timeline of the entire assessment procedure, including data collection, data analysis, assessment report writing, graduate programs committee review, CSSE graduate faculty approval, and final report submission. The procedure starts at the beginning of May, and the annual assessment report is submitted in the third week of June. The detailed milestones, deadlines, and description of each activity are tabulated in the table below.

A screenshot of a program

Description automatically generated with low confidence

It is worth mentioning that all the MS-DSE faculty members participated in the assessment procedure, including data collection. The MS-DSE curriculum committee thoroughly reviewed this report before the CSSE and DMS graduate faculty approved the assessment report. To facilitate efficient communication, we provide the graduate faculty with a summary of the key findings from the report.

Use of Results

11. Purposeful Reflection

11.1 A Consistent and Ongoing Process

The MS-DSE program committee consists of eight professors: Dr. Ku, Dr. Karmaker, Dr. Zhou, Dr. Qin, Dr. Billor, Dr. Ceyhan, Dr. Molinari, and Dr. Zheng, who are committee to arrange a minimum of two meetings per year to evaluate and improve the master’s degree program. The recommendations made by the MS-DSE committee will be further reviewed by the CSSE and DMS graduate faculty. In some cases, departmental policies will be forged to streamline the management of the MS-DSE program and master’s students. The graduate faculty always offer insightful feedback to the graduate programs committee to brush up on proposed policies and strategies. In case of disagreements occur among graduate faculty members, the department chairs – Dr. Narayanan and Dr. Abebe – will spearhead a solution as tradeoffs for the disagreements.

11.2 Implementation of the Above Process

The CSSE and DMS graduate faculty will continue improving both the curriculum as well as the opportunities for extracurricular activities. We look forward to future faculty development accompanied by the constructive and insightful feedback offered by the Office of Academic Insight on this MS-DSE graduate program annual assessment report.

The MS-DSE program committee review the assessment data in the second week of June, 2023 to analyze the findings observed from the assessment data. After the committee identified a handful of challenges, the committee proposed a solid action plan to tackle the obstacles preventing MS-DSE students from achieving exemplary academic performance. The detailed action plan, originated from the committee meeting, is documented in Section 12 below.

11.3 STAT 7940 and COMP 7980 Enrollment Analysis

We would like to share information on Fall 23- STAT 7940 capstone project enrollments. We allow students to take STAT 7940 or COMP 7980 after they complete the majority (at least 10) required courses. We have seven students registered for the STAT7940 Capstone project in the fall 22 semester. There are a total of five Data Engineering students completing the COMP 7980 in the summer 2022, fall 2022, and spring 2023 semesters. It means that these 12 students will most likely graduate by either the summer 2023 or the end of fall 2023 semesters.

The DSE committee will be working with two companies on this course: GE Aviation (they are committed to collaborating this year as well) and Southern Power as the second industry partner. These two industry partners are already committed to collaborating with the DMS DS Teams in the fall 23 semester.

12. Action Plan

**Action 1.** In this assessment cycle, the data is manually collected by CSSE and DMS faculty members. It takes approximately a neighborhood of 30 minutes to harvest data for courses with a large enrollment. This time-consuming approach, nevertheless, furnishes the rich assessment data allowing us to complete Section 8 on page 9.

As the first action item, we intend to make use of Canvas to align course components, including assignments and projects, to the MS-DSE program SLOs. It is practical to take full advantage of the Canvas system to automatically gauge assessment data. With regard to data collection, we have initiated a collaboration with Dr. Dorothy Thompson at the Academic Assessment Office to pilot an automatic way of gleaning data from the Canvas system. We had a kickoff meeting with Dr. Thompson on June 15. A second meeting with Dr. Thompson is scheduled in the first two weeks of the Fall’22 semester. A challenge to be addressed in the pilot study is how to deal with a common case where students from two or more graduate programs are taking the same graduate course managed by Canvas. This challenge is non-trivial because Canvas has no information about students’ graduate programs. We will collaborate with Mr. Paul Springfield – the Canvas system administrator – to dig out a solution that streamlines the data collection process with automation.

**Action 2.** Recognizing that the SLO 2 and SLO 4’s final scores are significantly lower than the other two SLO measures, the program committee plans to explore new ideas to foster students’ capability of utilizing practical computing technologies to implement data-driven solutions. In this joint MS-DSE program, the students have superior data science skills compared to their data engineering skills. We will pay heed to a handful of elective courses, in which the faculty members are encouraged to strengthen the software engineering skills of the MS-DSE students.

We are hopeful that SLO 2 and SLO 4’s overall score will be on par with those of SLO 1 and SLO 3. To achieve this goal, we will be focusing on a list of courses that emphasize SLO 2 and SLO 4 in the curriculum map articulated in Section 4 on page 3-5. These courses include, but not limited to,

* STAT 7000 Experimental Statistics I
* STAT 7940 Capstone Project
* COMP 6120 Database Systems I
* COMP 6130 Data Mining
* COMP 6710 Software Quality Assurance
* COMP 7980 Capstone Engineering Projects

**Action 3.** In the next assessment cycle, we will use of Qualtrics survey to collect data from COMP 7980 capstone engineering projects and STAT 7940 capstone projects. The COMP 7980 and STAT 7940 courses are intended to improve the communication skills of the MS-DSE students, who will deliver written and oral presentations to non-technical and technical audiences during the design and implementation phases of data science and engineering projects.

STAT 7940 Capstone Projects and COMP 7980 Capstone Engineering Projects have different teaching modalities: STAT 7940 is a course offered to a group of MS-DSE (Data Science ) students, whereas COMP 7980 is offered in the one-student-to-one-professor fashion. When the MS-DSE (Data Engineering) student enrollment climbs, we will cross list COMP 7980 and STAT 7940 so that both data science students and data engineering students can collaborate with one another developing capstone projects.

**Action 4.** The DSE committee will work closely with Dr. Zhou – the instructor for COMP 6130 Data Mining – to pilot new approaches to strengthening the programming skills of non-data-engineering students. We proposed the following ideas to bolster the student programming skillsets. Likewise, Dr. Zheng will continue implementing the strategies below to improve the academic performance of the graduate students who have week statistics background.

* Offer multiple difficulty levels for programming projects.
* Allow students to pick their preferred programming languages.
* Assist students to find group members to develop collaborative course projects.

**Action 5.** The DSE committee will schedule a meeting in the first week of June during the next assessment cycle to discuss various ways to strengthen the MS-DSE program. The discussion items on the meeting agenda are listed below.

* Exchange ideas on how we can make the program grow.
* How we improve the SLO2 and SLO4?
* Enrollment projections.
* The status of the proposed GCRT-DSE graduate certificate program.
* The accelerated bachelor’s/master’s program in the DMS and CSSE departments.
* Ways to enrich collaborations with other academic units/centers via the MS Data Science and Engineering Program.
* How to reinforce the collaborations with our industry partners?