

Data Science MS

The mission of the Master of Science in Data Science program is to train and develop world class data scientists.

Mission Statement (Full Description):

The Master of Science in Data Science (MSDS) is an online program designed for current and aspiring data science professionals looking to gain the advanced skills needed to manage, analyze, mine and understand complex data to make strategic decisions in their organizations. Through a combination of interactive coursework, collaborative group activities and online face-to-face classes, students gain the technical, analytical and communication skills needed to make meaningful data-driven decisions across various industries. The program aligns with SMU's mission of "...impart[ing] knowledge through teaching research, and service, while shaping individuals to contribute to their communities and excel in their professions in an emerging global society."

Does your program offer courses at an off-campus instructional site (not at SMU Dallas campus)?:

No

Does your program offer courses through distance education technology (e.g., asynchronous, synchronous, or both)?:

Yes

During which academic year were students first enrolled in this program?:

Prior to AY2023-2024

Progress:

Complete

Analysis Skills

SLO Step 1A: SLO Number:

1

SLO Step 1C: SLO Statement (Full Description):

Students will implement methods for data gathering, data cleaning, modeling, and deployment to generate relevant and reproducible conclusions [by the end of the program].

SLO Step 2A: Measure:

DS 6371 Final Exam: This final exam is a cumulative exam that covers data cleaning, data transformation, modeling, visualization, and communication (interpretation). Given that the problems on the exam depend on each of these skills, scoring a 70% on the Final Exam is a proxy for proficiency in each of the sub categories above.

SLO Step 2B: Type of Measure (check all that apply):

Objective Quiz or Exam

SLO Step 2C: Is Measure direct or indirect?:

Direct

SLO Step 3A: Target for Measure:

In AY23-24 approximately 70% of students scored a 70 or better on the final exam. Therefore, our measure will be to meet or exceed 70% of students scoring 70% or better on the final exam.

SLO Step 4A: Was the target met for this Measure?:

Met

SLO Step 4B: Results and Findings for this Measure:

Yes, in the Spring 2025 course, 87.5% (7 out of 8) students scored 70% or better on the Final Exam which is attached.

Attached Files

[OceanBacteria.csv](#)

[2025 Final Exam Statistical Foundations for Data Science - 401.csv](#)

[DS 6371 Final Exam Spring 2025.docx](#)

SLO Step 4C: Interpretation of Results:

(1). Interpretation of Results:

What are the strengths and weaknesses of student learning in this area?

Students showed strength in conducting hypothesis tests and performing linear regression. Weaknesses continued to be in interpretation and proper use of vocabulary.

How do results compare to previous years and between cohorts?

These results are a marked improvement over previous cohorts where the averages have been in the low to mid 70s. I believe ChatGPT may be related to this uptick.

(2). Documentation of Modifications or Previous Changes:

What actions or changes were made this year or in previous years to improve this student learning outcome?

We have doubled down on our efforts to do more live examples and student worked problems in live session breakouts. Additionally, we implemented TA office hours in addition to the professor office hours a number of years back which appears to continue to help. Finally, the oral exam seems to be very positively correlated with performance on the final exam. The thought being that students have studied the 23 oral exam questions intently over the semester which has helped their learning over the semester and reduced "cramming" and increased "learning" over the course of the semester.

What information and discussion prompted these actions?

Discussion with our Stat 2 professor, Dr. Jacob Turner, has helped quite a bit in developing and implementing these actions. He is also involved in delivering the oral exam which has helped shine line on its effectiveness as well.

How and when will impact of these changes be assessed?

We will continue to monitor the performance on both the final exam and oral exam to measure the impact of these interventions and initiatives.

SLO Step 5A: Use of Results for Seeking Improvement (Action Plan):

(1). How were assessment results shared with program stakeholders (e.g., program leadership, faculty, student, etc.)?

These results are shared at the program faculty meeting and there is subsequent discussion around how to improve and what actions need to be taken to do so.

(2). How has the program used assessment results for improvement? Using the assessment results, were there any decisions or changes made in curriculum, pedagogy, and other aspects that impact learning?

This is how the oral exam first came about. It was an idea that was brought up as a result of first poor performance and then better than expected performance on the final exam. When the average was low, we used the oral exam to help prepare students for the written final exam and to have better command of the material overall. When the final exam average got dramatically better, the program leadership suggested that the oral exam could be used to make sure students know what is reflected in the final exam score.

(3). What action(s) has the program taken over the course of the assessment cycle, or will take in the upcoming assessment cycle, for improvement in student learning?

A. Oral Exam Preparation in class ... students ask each other oral exam questions in breakout. In the oral exam, the student gets to pick a question and the professors pick a question. This semester we will have the students pick a question to ask another student in breakout during class. They will become the expert in this question and its response thus leading to better understanding overall and thus better performance on the final exam.

B. Per another faculty suggestion, I have added some additional practice exam to our GitHub page for independent practice.

C. We have continued with the alumni tutoring program.

SLO Step 5B: Type of Action:

Additional emphasis or time on content, Additional activities or assignments, Program leadership involvement, Faculty involvement

SLO Step 5C: Dialogue Participants (check all that apply):

Administrator, Faculty

SLO Step 5D: Evidence of Dialogue:

We presented these results in the MSDS faculty meeting on July 30, 2025. Slide 22 of the attached pptx.

Attached Files

[Faculty Meeting July 2025.pptx](#)

SLO Step 5E: Type of other Improvements (check all that apply):

Other

SLO Step 5F: Other Improvements (Full Description):

We continue to improve underline the importance of the oral exam which is to motivate to students to pay close attention to the material throughout the course async, assignments, HW and live session. This added attention to studies should be correlated with higher final exam scores which should be correlated with greater understanding and command of the statistical methods in this course.

SLO Step 6A: Status Update on Action(s) Identified in the Previous Assessment Cycle (Full Description):

Oral Exam:

The oral exam is referenced in every class and we implemented a break out where students ask each other questions from the list.

Alumni Mentoring program:

We continued this program this year with one alumni member helping one student in DS 6371. We are considering offering a badge or award to alumni that help current students.

MSDS Speaker Series:

We are running a third season of the MSDS Speaker Series aimed at increasing better analytical skills and application of those skills overall.

SLO Step 6B: Status Update on Previously Identified Action Plan(s):

In progress

Progress:

Complete

Ethics

SLO Step 1A: SLO Number:

2

SLO Step 1C: SLO Statement (Full Description):

Students will be able to critically evaluate and identify ethical issues related to data privacy, bias, and the responsible use of data while working on data science projects by the end of the program. Students will demonstrate their proficiency through case studies, project analyses, and written reflections, applying ethical frameworks to real-world data science scenarios.

SLO Step 2A: Measure:

Completion of the Ethics Portion of the Capstone A Project.

In the documentation section below you will find the description of the Capstone Project (CapstoneProjectTopics.pdf) and will find on page 1 paragraph 2, the description of the ethics portion of the paper. We have provided an additional description in PLO 2 Ethics Description.pdf also in the documents section below. Each paper will be graded as credit or no credit for the ethics portion of the paper. To get a "credit", the student must address the ethical considerations his or her project represents and how they have been mitigated or at least how they were handled in the project and/or in practice. In addition, for reference, in the documentation section there is also an example of a "good" ethics section that was given "credit" (A Personalized Approach to Understanding Human Emotions.pdf). Finally, we have included an Excel file of scores from the Spring 2023 Capstone A class (PLO2 Spring 2023.xlsx).

Attached Files

[CapstoneProjectTopics.pdf](#)

[PLO 2 Ethics Description \(1\).pdf](#)

[A Personalized Approach to Understanding Human Emotions.pdf](#)

[PLO Spring 2025 .xlsx](#)

SLO Step 2B: Type of Measure (check all that apply):

Capstone project

SLO Step 2C: Is Measure direct or indirect?:

Direct

SLO Step 3A: Target for Measure:

Given the importance of ethics in data science and the fact that every student should be able to complete the measurement if they put forth the thought and effort, we aim for 100% of students to receive "credit" for the ethics portion of their Capstone A (DS 6210) paper.

SLO Step 4A: Was the target met for this Measure?:

Met

SLO Step 4B: Results and Findings for this Measure:

As seen in PLO Spring 2025.xlsx, 100% of the students in the Spring 2025 course received "credit" for the ethics requirement in DS 6210.

Attached Files

[PLO Spring 2025 \(1\).xlsx](#)

SLO Step 4C: Interpretation of Results:

This outcome confirms that students were effectively prepared to engage with ethical dimensions of data science by the end of the program. The scaffolding built into earlier coursework, along with clear expectations in DS 6210, appear to be successfully supporting student proficiency. The integration of ethics into the capstone also ensured authenticity, encouraging students to apply frameworks in a real-world, project-specific context.

SLO Step 5A: Use of Results for Seeking Improvement (Action Plan):

While the 100% achievement rate demonstrates strong program performance, the following steps will be taken to ensure continued success and deepen ethical engagement:

- **Maintain clarity and support** around the ethics requirement in DS 6210, including rubrics and example analyses.
- **Reinforce ethical reasoning earlier** in the curriculum by embedding brief ethical scenarios or reflection prompts in prerequisite courses.

SLO Step 5B: Type of Action:

Additional activities or assignments,Redesign of activities or assignments

SLO Step 5C: Dialogue Participants (check all that apply):

Faculty

SLO Step 5D: Evidence of Dialogue:

I have emailed our Capstone professor to discuss the implementation of the action plan. I have attached the email in *Assessment Action Plan.msg*.

Attached Files

[Assessment Action Plan .msg](#)

SLO Step 5E: Type of other Improvements (check all that apply):

Other

SLO Step 5F: Other Improvements (Full Description):

We changed the course from a two semester course (one unit a course) to a one semester course and made it a 2 hour course. We believe this will create more fluid and complete papers due to the consistency in which the students work on them (less down time that we have observed in the past.)

SLO Step 6A: Status Update on Action(s) Identified in the Previous Assessment Cycle (Full Description):

Action from last cycle: We implemented the MSDS Speaker Series this year and each season (including this past season) will have a speaker that talks specifically to data ethics. `

Update: We are planning to have a speaker in season III of this years speaker series from Humana to speak on ethics. Dr. Cheun will also have a guest speaker on ethics in the fall and she had one in the Summer.

SLO Step 6B: Status Update on Previously Identified Action Plan(s):

In progress

Progress:

Complete

Communication

SLO Step 1A: SLO Number:

3

SLO Step 1C: SLO Statement (Full Description):

Students will develop the ability to communicate complex data findings clearly to non-technical stakeholders [by the end of the program].

SLO Step 2A: Measure:

At least 80% of students will score “proficient” or higher on a data presentation assignment (Case Study 2) that requires them to translate technical data insights into actionable strategies for business._

SLO Step 2B: Type of Measure (check all that apply):

Presentation

SLO Step 2C: Is Measure direct or indirect?:

Direct

SLO Step 3A: Target for Measure:

We aim to have 80% of students score above a 90%. The benchmark of 80% is close to percent that achieved a 90% in AY22-23.

SLO Step 4A: Was the target met for this Measure?:

Met

SLO Step 4B: Results and Findings for this Measure:

100% of students scored 90% (10/10) or better.

SLO Step 4C: Interpretation of Results:

We were very pleased with the performance of the students on this exam. A potential weakness of this measurement is the potential use of generative AI that could be contributing to the high performance. However, this is a recording of a live presentation and students are checked for signs of reading in their presentation or signs of code generation in their github repos. These students still achieved high scores after these checks.

SLO Step 5A: Use of Results for Seeking Improvement (Action Plan):

Action Item 1: We will have the professor of our new Storytelling course to present to the class on presenting to the C-Suite and how presenting to the CEO is different from presenting to the CTO. While all presentation were above the threshold, the principal area of improvement was differentiating the delivery to the CEO and the CTO.

Action Item 2: We have introduced Retrieval Augmented Generation (RAG) into this course as well as deployment of an app using Hugging Face Spaces. We will continue polishing this content and will look to make the creation and deployment of the app a requirement for project 2.

SLO Step 5B: Type of Action:

Additional activities or assignments, Faculty involvement

SLO Step 5C: Dialogue Participants (check all that apply):

Faculty

SLO Step 5D: Evidence of Dialogue:

These results were shared at the faculty meeting on July 30, 2025. (Slide 24 of the attached pptx deck).

Attached Files

[Faculty Meeting July 2025.pptx](#)

SLO Step 5E: Type of other Improvements (check all that apply):

Other

SLO Step 5F: Other Improvements (Full Description):

The only major improvement planned right now is the addition of our Professor of the Storytelling with Data course to talk about presenting to CEOs versus CTOs.

SLO Step 6A: Status Update on Action(s) Identified in the Previous Assessment Cycle (Full Description):

We did not assess this SLO in the previous cycle; however, we continue to use the RShiny dashboards that are listed as action items in the AY22-23 cycle as well as the Jekyll pages introduced in the same year. To the later, students have really enjoyed and told us they have benefitted from having a nice polished website to create their portfolio.

SLO Step 6B: Status Update on Previously Identified Action Plan(s):

In progress

Progress:

Complete

Team Work

SLO Step 1A: SLO Number:

4

SLO Step 1C: SLO Statement (Full Description):

Students will develop the ability to collaborate effectively within diverse teams to solve complex data science problems [by the end of the program].

SLO Step 2A: Measure:

In *Doing Data Science (DS 6306)*, Case Study 1 is a group project composed of two group members. The students are responsible for cleaning, analyzing and creating inference from a dataset provided by the instructor. In addition, the students are responsible for communicating the results of this project by both presenting their results live in class as well as in a 5 minute recorded video recorded and stored on the student's personal YouTube account. The students are required to give a detailed review of their partner's contribution and performance (Scale of 1 (None) to 10 (Exemplary)) and the professor provides a rubric graded assessment of the written and orally communicated results. There are two measures below to assess (1) that the team was able to work together and (2) the quality of what they produced as a team.

SLO Step 2B: Type of Measure (check all that apply):

Presentation ,Rubric,Written paper/project

SLO Step 2C: Is Measure direct or indirect?:

Direct

SLO Step 3A: Target for Measure:

There are two Targets for PLO4:

- 1.We aim to have 80% of students rate (score) their partner an 8 or higher on their Peer Review in order to assess cooperation and effective interpersonal communication.
- 2.We aim to have 80% of students score a 90% or greater on the project to assess the product of their teamwork.

Both of these measures were based on approximate student performance from AY22-23.

SLO Step 4A: Was the target met for this Measure?:

Met

SLO Step 4B: Results and Findings for this Measure:

- 1.We aim to have 80% of students rate (score) their partner an 8 or higher on their Peer Review in order to assess cooperation and effective interpersonal communication.

Commented [BG1]: Rubric:

I. Teamwork & Collaboration (20 points)

- Peer Evaluation (10 pts):** Rating of partner's contribution on a scale from 1 (None) to 10 (Exemplary). Submitted confidentially by each student.
- Equitable Contribution (10 pts):** Evidence that both members contributed fairly to key components of the project. Instructor assesses based on observed work balance and student input.

II. Data Cleaning & Analysis (25 points)

- Data Cleaning & Preparation (10 pts):** Quality of data wrangling, handling of missing values, transformations, and preprocessing steps.
- Analytical Rigor (10 pts):** Appropriateness and correctness of applied statistical or machine learning techniques.
- Reproducibility (5 pts):** Code and documentation are clear, complete, and enable others to replicate results.

III. Inference & Insight (20 points)

- Clarity of Inference (10 pts):** Conclusions drawn from the analysis are clear, logical, and grounded in data.
- Actionable Recommendations (10 pts):** Insights are meaningful and translated into specific, realistic business or real-world strategies.

IV. Communication (35 points)

- Written Report (10 pts):** Well-organized, clear, and professionally formatted report. Includes appropriate visualizations and explanations.
- Live Presentation (10 pts):** Verbal delivery is clear, confident, and within time limits. Effective use of visuals and engagement during Q&A.
- Recorded Video (10 pts):** A concise, 5-minute summary of the project findings posted on YouTube. Must be clear, accurate, and well-paced.
- Visual Aids (5 pts):** Use of charts, tables, and slides enhances understanding across all formats (report, live presentation, and video).

In the Spring 2025 cohort, 80% of the student scored an 80% or higher. They were put in groups of two and one group had a pretty significant issue and they ended up breaking up and doing the project on their own.

2. We aim to have 80% of students score a 90% or greater on the project to assess the product of their teamwork.

Despite breaking up, the two students that did individual projects did well as did all the teams. 100% of students scored a 90% or better showing that for those that stayed in groups, not only did they appreciate and respect each others contributions, but they also did an objectively good job on the analysis and presentation. In short, in most cases, they were a good team.

See the attached document for the raw score reports.

Attached Files

[Spring 2025 DS 6306 Peer Review and Final Score.xlsx](#)

SLO Step 4C: Interpretation of Results:

Strengths in Student Learning:

Students in the Spring 2025 cohort demonstrated strong mastery in both the collaborative and analytical components of the project. Specifically:

- **Team Dynamics:** 80% of students rated their partners 8 or higher in peer evaluations, meeting the benchmark and indicating strong interpersonal communication, shared responsibility, and cooperation among most student teams.
- **Project Performance:** 100% of students scored 90% or higher on the final project, showing high proficiency in data cleaning, analysis, inference, and communication. This reflects students' ability to apply technical concepts and communicate findings effectively to both technical and non-technical audiences.

Even in the one instance of group conflict—where a pair separated and completed the project individually—both students performed exceptionally well. This suggests a level of individual preparedness and resilience, even in non-ideal collaborative settings.

Weaknesses in Student Learning:

- **Conflict Resolution and Team Resilience:** The occurrence of a team breakdown highlights the need for improved support mechanisms for conflict resolution and more structured check-ins to proactively surface and address issues within teams.

Comparison to Previous Years and Cohorts:

- Peer collaboration results (80% \geq 8 rating) are consistent with AY 2022–2023 benchmarks, indicating stable team effectiveness over time; although, we would like to see growth and functional teams.
- The increase to 100% of students achieving \geq 90% on the final project in Spring 2025 reflects an upward trend in student proficiency and is consistent with the changes made to the assignment and assessment criteria in prior years having a positive impact.

SLO Step 5A: Use of Results for Seeking Improvement (Action Plan):

Action Plan: To proactively support group collaboration and ensure peer evaluations provide meaningful, accurate feedback that reflects students' contributions.

1.
 - At the start of the project, we will have a breakout with each student where they fill out a calendar that has all their availability. We have found that more than 90% of the time, when a team breaks up, it usually started with a missed meeting or scheduling conflict.
 - Additionally, in this same breakout, we will have them complete a small team building activity like learning a bit about each other and then presenting what they learned about their partner to the class. :)

SLO Step 5B: Type of Action:

Additional activities or assignments

SLO Step 5C: Dialogue Participants (check all that apply):

Faculty, Student

SLO Step 5D: Evidence of Dialogue:

We presented these in the July 30 Faculty meeting. It is slide 25 of the attached

Attached Files

[Faculty Meeting July 2025.pptx](#)

SLO Step 5E: Type of other Improvements (check all that apply):

SLO Step 5F: Other Improvements (Full Description):

I think it would be a good idea to have one of our psychology professors or someone who specializes in teamwork come in to talk with the students about teamwork, communication and the value and power of empathy and cooperative thinking.

SLO Step 6A: Status Update on Action(s) Identified in the Previous Assessment Cycle (Full Description):

This PLO was not assessed in the last cycle.

SLO Step 6B: Status Update on Previously Identified Action Plan(s):

In progress

Progress:

Complete

Improve Alumni Involvement

PG Step 1A: PG Number:

1

PG Step 1C: PG Statement (Full Description):

One of the biggest facts that separates MSDS from other data science programs is the fact that over the last 10 years we have graduated over 800 alumni. I would like to increase their involvement and activity in the program by offering three specific initiatives:

A. Quarterly Alumni Meetings: These meetings feature a faculty member or other expert giving a talk on a current research topic. We are currently deep into year two but have not collected attendance data on these meetings. (I have attended each meeting and estimate attendance to average about 10.)

B. A new initiative is one of continuing / maintaining education: **MSDS Update!** Each month we would have a professor teach the current unit in their course for the alumni. This would be

from one of our required courses like DDS, Stat 1, Stat 2, ML 1 or QTW. This would be a low lift for faculty since they have already prepped and delivered that lesson recently and should be useful and fun for alumni since they are either refreshing or updating something they have studied in the past.

C. The final initiative is to reignite the push to get alumni access to the courses they have taken in the past. This was once an advertised benefit of signing up for the program but has always been difficult to get 2U to provide access. I have had several students inquire about this benefit over the years and I do not know of a single student who has been successful in securing access. 2U has always said that it was possible but that students had to "opt in" and specifically request the course they want access to each year. I believe this barrier to entry is intentional and aimed at reducing the number of students that actually use benefit. I would like to make this available to all students to where they don't need to opt in rather have access to the material by simply logging in. For what it is worth, UC Berkely (another 2U affiliate) still advertises this benefit as MIDS for Life: <https://ischoolonline.berkeley.edu/data-science/>

Below these will be referenced as A, B and C.

Value:

It is also important to note how keeping and increasing the involvement of the alumni helps the MSDS program. Many of our current and former students were referred by alumni members and maintaining or increasing involvement should at least maintain and maybe increase these referrals.

Additionally, with the accelerating evolution of AI and generative AI, it reflects well on the MSDS program and SMU in general to have 800 alumni members that on the cutting edge of technology and are thus in high demand in their individual industries. This also works in reverse to where if our alumni fall behind, this will could have a significant negative effect on the MSDS and SMU brand.

Finally, these events are open to both alumni and current students. Providing opportunities for current students to network with the alumni opens up hiring, learning and other opportunities to students currently in the program.

PG Step 2A: Measure:

A. The measure is simply the four quarter moving average at the quarterly alumni meetings.

B. The measure is simply the four month moving average at the monthly **MSDS Update!** program.

C. Activity can be tracked by 2U. We would track overall usage across all courses three times a year. Our measure would be the three trimester moving average.

PG Step 2B: Is Measure direct or indirect?:

Direct

PG Step 3A: Target for Measure:

A. The yearly target is to have the 4 quarter moving average that ends in (including) Quarter 2 (June) be higher than the previous year's four quarter moving average ending in (including) Quarter 2 (June).

B. The yearly target is to have the 4 month moving average that ends in (including) June be higher than the previous year's four quarter moving average ending in (including) June.

C. The yearly target is to have the 3 trimester moving average that ends trimester 1 (May) be higher than the previous year's 3 trimester moving average ending in (including) trimester 1 (May).

PG Step 4A: Was the target met for this Measure?:

Partially Met

PG Step 4B: Results and Findings for this Measure:

A and B were not met and C was met.

A. We discontinued the Alumni meetings due to lack of attendance and replaced it with the MSDS Speaker Series. We had 5 speakers last year which drew an average of 40 students and community members to each meeting.

B. We also discontinued the MSDS Update program in favor of MSDS Workshops. We have had one workshop that we capped at 20 and filled all 20 seats. We are moving to take the cap off for the next workshop. There are three scheduled for the Fall.

C. This one is still ongoing and we have added 2 alumni to the Stat course on going section. The moving average is now 4. This is the moving average of how many people are enrolled in the ongoing section rather than adding each trimester. This is a very popular program to discuss with students but not one that many students go out of their way to participate in or take advantage of.

PG Step 4C: Interpretation of Results:

Our assistant director, Dr. Jonathan Thompson, the faculty and I made the decisions above to discontinue and reimagine A and B above. We will measure the attendance of the MSDS Speaker Series and the Workshop series on a per event basis going forward.

For A, the MSDS Speaker Series, we will set the target to average 45 per event for Season 3.

For B, the MSDS Workshop Series, we will set the target to average 25 per workshop.

For C, the alumni addition to permanent audit courses, we will target to add one alumni member per semester.

PG Step 5A: Use of Results for Seeking Improvement (Action Plan):

Action plan:

The MSDS program will begin to post weekly about each event, A, B and C our linked in Page. Additionally, when possible, we will post the content on the our MSDS YouTube Chanel.

PG Step 5B: Dialogue Participants (check all that apply):

Staff

PG Step 5C: Evidence of Dialogue:

Please see the email .msg file below for evidence of this dialog.

Attached Files

[Posting about Speaker and Workshop Series.msg](#)

PG Step 5D: Type of other Improvements (check all that apply):

PG Step 5E: Other Improvements (Full Description):

PG Step 6A: Status Update on Action(s) Identified in the Previous Assessment Cycle (Full Description):

Ideas from previous cycles are now moot given the new direction described above.

PG Step 6B: Status Update on Previously Identified Action Plan(s):

Not applicable for this cycle (explain in Step 6A)

Progress:

Complete

