

Graduate Council
Questions for annual online degree program check-in
Master of Information and Data Science (MIDS) 2016

The following questions will form the basis of the four-year review of online degree programs. As a way of checking in on how your program is doing before that review, please answer these questions based on the past year.

Wherever possible, please provide statistics or evaluation comments to help the GC gauge the program's development. The report should not exceed 10 pages.

- 1. What is the quality of the admitted students (e.g., test scores, GPA) compared to on-campus degrees offered by your unit, or peer programs at other institutions (if known)? How does the diversity of your admitted students compare to similar on-campus and peer institution programs? What percentage of students are expected to graduate on time? What is the attrition rate?**

The quality of students admitted to the online Master of Information and Data Science (MIDS) program remained high in 2016, with an average GPA of 3.49 and strong average standardized test scores, which are required for admission. For 2016, average admitted student test score percentiles were the following: GRE Quant - 83rd percentile, GRE Verbal - 84th percentile.

The MIDS program continues to focus on increasing the number of women in the field of data science. We are pleased to be making progress on this front. In 2016, 27% of admitted students were female, compared to 21% in 2015.

The program also attracts students of all ages; in 2016 the average admitted student age was 34, while the youngest admitted student was 21 and the oldest was 65.

The quality of students admitted to MIDS compares favorably to those admitted to our on-campus Master of Information Management and Systems (MIMS) program. Average undergraduate CGPAs are similar: 3.49 MIDS v. 3.58 MIMS, as are GRE quant scores: 83% MIDS v. 86% MIMS and GRE verbal scores: 84% MIDS v. 81% MIMS

MIDS 2016 admitted students (n=300)

Avg GRE verbal percentage: 84

Avg GRE quant percentage: 83

Avg domestic UG CGPA: 3.49

% women: 27

Countries represented: 15

MIMS 2017 admitted students (n=75)

Avg GRE verbal percentage: 81

Avg GRE quant percentage: 86

Avg domestic UG CGPA: 3.58

% women: 47

Countries represented: 13

Self-reported ethnicity for 2016 MIDS admits

Ethnicity	#	%
2 or more	23	8
Asian or Asian American	34	11
Black or African American	4	1
Chinese/Chinese-American	53	18
Hispanic or Latino	6	2
Other	7	2
Other Asian/Asian American	26	9
Unknown	15	5
White or Caucasian	132	44

It is also noteworthy that the total number of underrepresented minority admitted MIDS students was 24, or 8%. This was substantially more than our MIMS admitted student pool, which totaled 3, or 4%.

From a retention standpoint, 91% of MIDS students are expected to graduate on time. The program has had a 5% attrition rate inception to date.

2. To what extent has the degree met its academic objectives in the past year as laid out in the initial proposal? What metrics have been used to measure these outcomes and how have the evaluations been done?

The academic objectives we set for this program were to address the practical challenges facing managers and other professionals working with large, messy, often incomplete datasets; to open access to working professionals in California and elsewhere who are interested in education that equips them to insightfully analyze the floods of new data to solve problems and to achieve organizational and societal goals; to provide hands-on practice working with complex, unstructured and user-generated data to identify new ways to inform decision-making; to integrate insights from social science and policy research, as well as statistics, computer science, and engineering, in an authentically multidisciplinary program; and to set the benchmark for high quality, online professional education of data scientists.

The multidisciplinary breadth of the MIDS curriculum is reflected in the fact that three of our courses were developed by, and are taught by, social scientists: Data Science W201, Research Design and Applications for Data and Analysis, Data Science W231, Behind the Data: Humans and Values, and Data Science W341, Experiments and Causal Inference. These complement more technical courses such as W203, Statistics for Data Science, W205, Storing and Retrieving Data, W207, Applied Machine Learning. We also offer W209, Data Visualization and Communication, which is a hybrid course.

In the past year, we have successfully achieved our academic objectives regarding practical applications of multi disciplinary techniques to data science challenges. We now offer thirteen courses, the majority of which are project-based and require students to work directly with publicly-available datasets in real-world professional scenarios. Our newest formal course, Python Data Science W200: Python Fundamentals for Data Science, which was approved by COCI for its first offering as a credit-bearing class, not only introduces the fundamental computer science skills needed for the field (in this case, Python programming), but also engages students in applying those skills to data analysis, bridging the gap between theory and practice.

Many of our foundational classes continue to integrate theoretical content with intuition and practical examples. Data Science W207: Applied Machine Learning incorporates probability, statistics, and linear

algebra, while also teaching students to apply machine learning techniques directly to current problems in the field, and Data Science W266: Natural Language Processing introduces linguistics and advanced machine learning techniques with a focus on practical applications such as information extraction, machine translation, sentiment analysis, and summarization.

We have also met our goal of increasing access to data science education for working professionals in California and beyond. There were over 400 students currently enrolled in the MIDS program at the end of 2016, of whom over 150 live and work in California, approximately 200 live and work elsewhere in the United States, and more than 50 live and work outside the United States. We now have over 200 alumni worldwide as well. Our students and alumni are forming a robust network of data science professionals equipped to address new information challenges for organizations across the globe.

Course evaluation data

Students fill out course evaluations two times over the course of each term. The first set of evaluations are undertaken during the 5-6th weeks of instruction and are made available to instructors immediately in order to facilitate improvements to instruction. These evaluations are largely qualitative and are used to assess not only the quality of instruction, but also the quality and timeliness of instructor feedback and the ease of use of the technology. Final course evaluations are conducted during the last two weeks of instruction and are used to assess instructor and course effectiveness as well as the clarity and usefulness of course materials. Both quantitative and qualitative feedback is solicited and used in course development and instructor training.

The average rankings for our Spring 2017 courses were 5.82/7 in Instructor Effectiveness and 5.55/7 in Course Effectiveness. Five out of thirteen courses were ranked at 6 or better in Course Effectiveness. All but four courses have seen improvements in ratings for Instructor Effectiveness and all but five have seen improvements in ratings for Course Effectiveness.

Student feedback is assessed by program faculty and staff and channeled into improvements to course materials and structure; and into training and best practices for instructors.

Job placement data

The most important outcome measure for any professional degree is placement of graduates. We are delighted with the updated career outcomes information for MIDS, which is drawn from an outcomes survey that is sent to all graduates upon completion of the program. The data below is from 2014, when we launched the program, through 2016 and includes 129 graduates.

Median Salary: \$120,000

Median Annual Bonus: \$15,000

Graduates reporting a salary increase: 72%

Graduates reporting a promotion: 30%

Graduates reporting a new employer: 58%

*Top employers of MIDS graduates (# employed)**

Amazon (6)

Google (5)

Capital One (4)

Apple (3)

eBay (3)

Facebook (3)

IBM (3)

*Job titles (graduates with title as % of total graduates)**

Data Scientist (including: Principal, Senior, Lead, Staff, Associate) **29%**
Engineer/Architect/Developer (includes Data, Business Intelligence, Machine Learning, Software) **21%**
Analyst (including: Data, Product, Business) **9%**
Manager - Data, Analytics, Engineering **7%**
Director - Data Science, Analytics **7%**
C-level/VP-level **5%**
Founder/Co-founder **4%**
Consultant **3%**
Product Manager **3%**
Scientist/Researcher **2%**
Quant/Trader **2%**
Other/unknown **6%**

* Employer and job title information is based on outcomes surveys and LinkedIn analysis

Ladder faculty oversight

We regularly share the MIDS academic and career outcomes with the ladder faculty in the I School, as well as those who are actively teaching in the program. The faculty appear to be quite happy with the progress of the program overall.

3. What changes, if any, have been made in the delivery of the degree, either for individual courses or for integrating components of the degree (e.g., developing a community, the capstone project, advising)? What changes are anticipated in the upcoming year?

We made many changes to MIDS in response to student, faculty, and staff feedback in 2016, including

- The School provided an institutional subscription to *Slack* (a tool for real-time messaging, archiving, and search for teams) for all I School students. This is used in many of our MIDS classes, with faculty communicating with students, as well as among subgroups of students, e.g. those who live in Seattle, as well as between the I School and the entire community. It has made a huge improvement in the sense of community among MIDS students, and between MIDS and MIMS students.
- We launched one new advanced elective course: *Natural Language Processing*. It is now our most popular elective. The *Python for Data Science* class we reviewed by COCI and introduced as an elective rather than just an optional bridge class.
- We doubled the size of our Academic Advising/Student Affairs staff, from 2 to 4
- We established a pathway for communication of career information between faculty and career services staff, allowing for shared understanding of students' professional goals and strategies.
- We offered a webinar series on Deep Learning, featuring MIDS faculty as well as local area experts in the field, in response to student demand for more deep learning content.
- We launched a series of alumni panels showcasing data science careers for existing students.
- The MIDS students hosted a Data Science Challenge for MIMS students

- We hosted MIDS meet ups for prospective and current students as well as alumni in areas where there is a strong geographic concentration of students, including New York, Texas (Houston), and Washington D.C.
- Our staff continued their attendance at diversity conferences for recruitment and outreach. In 2016 this included the Grace Hopper Celebration of Women in Computing and the Chick Tech conference.

We are updating and/or rebuilding several of our courses in 2017, largely in response to student feedback: We are updating W209, Data Visualization and Communication to include more hands-on D3 and java script exercises, and W210 Research Design for Applications of Data and Analysis to include more current examples of using data to solve organizational problems, as well as to improve some of the short lectures. Finally we are engaged in a complete rebuild of W205, Storing and Retrieving Data, so that it has a clear narrative arc following the evolution of the tools (from SQL to Hadoop to Sparc).

We will be offering the MIDS class, W241 Experiments and Causal Inference, in a hybrid format to on-campus students. We are also planning several regional meet ups of prospective and current MIDS students as well as alumni for 2017, in the South Bay, Denver, and Boston. We have also increased our sponsorship level at the Grace Hopper Conference.

4. Have there been changes in the competitive environment (e.g., new programs at peer institutions) that have changed the landscape for this degree?

Many new online and on campus data science/analytics Master's programs were launched in 2016, several from peer institutions. Some notable new programs are listed below. In addition, we have seen new courses offered to data science students including Natural Language Processing and Advanced Statistical Modeling. Media coverage of data science and related fields continues to evolve away from the term "Big Data" and towards more specific skills and tools such as "Predictive Analytics", "Deep Learning" and "Machine Learning" or "R" and "Python."

We are delighted that in spite of the growing competition, MIDS was recently ranked as one of the top six data science Master's degrees in the US, along with: CMU, MIT, Stanford, Northwestern, and NYU.

Notable new data science programs:

In August 2017, Georgia Tech will launch their online Master of Science in Analytics, which features an interdisciplinary curriculum drawn from Georgia Tech's College of Engineering, College of Computing and the Scheller College of Business. The program currently offers two focused tracks in "big data" and "analytical tools," with a third "business analytics" track in the works.

August 2017 will also see the launch of Notre Dame's online Master of Science in Data Science program. In collaboration with AT&T and offered by the Department of Applied and Computational Mathematics and Statistics, the Department of Computer Science and Engineering, the Mendoza College of Business and the Department of Psychology, this degree program will prepare graduates for careers as data scientists in a wide range of industry fields including management, marketing, information technology, government policy, health care, finance, education and scientific research.

The Master of Science in Data Science will launch in September 2018 at Harvard University. Developed by the Harvard computer science and statistics faculty with input from industry experts, the program will offer preparation in statistical modeling, machine learning, optimization, management and analysis of massive data sets, and data acquisition. Program requirements include a technical core of four courses in data science, computer science, and statistics, as well as a capstone research project that partners students with academia, government, finance, technology and health care companies.

Despite these new, competitive programs entering the market, we continue to see record demand for MIDS. We have built competitive advantage in the landscape by designing a professional program with a focus on educating data science leaders via a multi-disciplinary curriculum that draws insights from the social sciences, computer science, statistics, management, and law.

5. What is the degree of student satisfaction in the advising and community-building aspects of the degree program?

Student satisfaction with academic advising is evaluated in MIDS using a net promoter score, which is an index that ranges from -100 to 100 and measures the willingness of students to recommend the academic advising team to their peers. Net promoter score is a high bar with which to measure satisfaction. A world-class score is anything above a 50. *In 2016 the MIDS academic advising team received an overall net promoter score of 68.*

We have also observed students developing a community within the MIDS program. The functionality of our online learning platform allows for students to form virtual groups. Students have used this feature to launch groups related to both their academic/professional endeavors and social interests. In 2016, a program survey indicated that 72% of students agreed with the statement, “I feel like a member of my university community” and 83% of students agreed with the statement, “This program has helped me develop a network with fellow students.” We are proud that in Dec, 2016 we had 129 MIDS graduates.

6. What has been the educational benefit of the capstone project or comprehensive exam?

The capstone project is an essential, and highly valued, component of the MIDS degree program. The students are required to enroll in a workshop course (W210) that brings them together in small teams to design and complete a full data science project using openly available data sets. The project provides students with the opportunity to integrate all of the core data science skills and concepts they have learned throughout the program—and prepares them for professional success in the field by focusing on the ability to collaboratively develop and communicate their work in both written and oral form. We assign two experienced instructors to teach this class: one with business, the other with data science skills.

Students present the projects in an online Capstone Project Showcase for the entire MIDS community, which is held at the end of each of three semesters. At that time we select a winner of the Hal R. Varian MIDS Capstone Award for the best capstone project. The quality of the projects has been high. For example, one team was invited by an instructor to present their project at Google. The project seeks to mitigate online gender harassment through (1.) user feedback, (2.) empathic innovation, and (3.) data science products. This has led to an informal partnership between the MIDS alumni and the lead scientist for a Google Jigsaw project, a joint experiment and publication, and ongoing collaborations.

One graduate, writing with advice to MIDS colleagues about the job search process, underscored the value of the capstone project in the job interview process. He advised them to make a personal website and highlight their capstone project: “I linked to Github and my project websites whenever possible. I made the first project on the page my capstone project. People responded well to my capstone, about machine learning in baseball. What I didn’t realize is that people almost always only read the first project. So do your best to have one great project, put it first, expect people to read about it . . .this makes the capstone project even more important, since it’s a natural choice for that first project.”

Students consistently rank W210, the Synthetic Capstone course, above the other courses in the program for both course and instructor effectiveness. Some of the written comments include:

- This was a great way to end an amazing MIDS program!

- First and foremost, the course is designed (as it should be) to encompass all aspects of the MIDS program. It builds on all material and offers a major contribution to any graduate's work portfolio. Students have leeway and flexibility to work on projects that they find to be intellectually stimulating, which is usually half the battle in projects of this magnitude. Also, the course is designed around the major project, which helps students move in the right direction throughout the term.
- Personally, I cannot thank both [of you] enough for guiding us through a wonderful semester of learning. Thanks to the way you present the course, your constant encouragement and guidance and my two wonderful teammates, I truly enjoyed the course in the end. It may have concluded my formal education in MIDS, I believe I will benefit from this experience for years to come.

7. Has offering the online degree program affected any on-campus programs (positively or negatively)? If so, how?

The MIDS program has raised the visibility of the I School, which has had positive impacts for both of our on-campus programs, MIMS and the PhD. It has grown our applicant pools, attracted new recruiters to MIMS as well as MIDS, and generated the revenue to build a career services staff.

In 2017 we plan to offer the MIDS class *W241, Experiments and Causal Inference* to on campus students in a hybrid format (students view videos and attend live sessions.) If that succeeds we plan to expand the hybrid offerings of MIDS courses not only for MIMS students, but also students elsewhere on campus.

8. Have any financial concerns emerged in the past year related to the development or delivery of the degree?

For the first three years of the program, we covered all of the costs for MIDS students attending the on-campus Immersion program. However the cost of hotel rooms, food, and other services have increased faster than we anticipated; we are now spending approximately \$2000 per student. We have decided to shift more of these costs to the students. We will ask them to pay for their hotel room and a \$500 Immersion fee. Many, but not all, students can have these costs defrayed by their employers. We will also provide need-based support for students who will have trouble covering these costs.

9. How and to what extent have campus-wide resources been drawn into the online program? Have there been any benefits for students in the online program?

MIDS students make very little use of campus resources, aside from online access to the campus library. They typically spend only one week on campus for the Immersion program. They get Cal ID cards and they purchase sweatshirts, but they do not avail themselves of most campus student services (athletics, tech support, ASUC, and so forth) and they do not have access to health services.