

Contact: Rachel Boon

**REQUEST FOR NEW PROGRAM AT THE UNIVERSITY OF IOWA: MASTER OF SCIENCE
IN DATA SCIENCE**

Action Requested: Consider approval of the request by the University of Iowa for a Master of Science in Data Science in the Graduate College.

The Council of Provosts and Board office support approval of this program.

Description of proposed program. The Department of Statistics and Actuarial Science is proposing a graduate program leading to a Master of Science in Data Science (MSDS) in the Graduate College at the University of Iowa. The MSDS will require a minimum of 30 semester hours of coursework for graduation. The coursework includes 6 core courses covering the fundamentals of data science including (1) foundations of probability and statistics; (2) data storage, access and management; and (3) data visualization, exploration, modeling, analysis and uncertainty quantification. Based on their interests and career goals, students also pick three electives (9 hours) from a wide variety of courses on specialized data science topics offered by the Departments of Statistics, Biostatistics, Computer Science and Business Analytics to enhance their skill set. Additionally, through a required capstone project, students will acquire hands-on experience in solving real-world problems that will include an emphasis on communication skills and ethics related to working with and learning from data.

The proposed MSDS program will be mutually beneficial to our strong undergraduate and graduate programs in actuarial science, including the BS in data science. The Department of Statistics and Actuarial Science has been designated as a Center of Actuarial Excellence (CAE) since 2009; it is a distinction shared by only 17 major research universities worldwide. The insurance industry has been expanding the uses of data in directing their business; consequently, predictive analytics emerges as one of the predominant topics in professional examinations and actuarial science research.

Academic objectives. The goal of the proposed MSDS program is to train data scientists who have the analytical and technical skills to explore, formulate and solve complex data-driven problems in science, industry, business and government.

- The proposed program focuses on the theory, methodology, application and ethics for working with and learning from data.
- With emphasis on statistical thinking and foundations, the MSDS program will rigorously prepare its graduates with the abilities to develop and implement new or special-purpose analysis and visualization tools.
- The program will promote a fundamental understanding of how to quantify uncertainty when making data-driven decisions.

The program will prepare graduates interested in academic, industry, business or government positions that involve data visualization and modeling, managing reproducible data analysis workflows, and collaborating/communicating with scientists and other data stakeholders.

Need for program. Data of various types, complexity and sizes are increasingly collected and used across fields, widening the integration of data directed applications in society, science and business. Companies and organizations of every size and industry – from Google, Amazon, medical research centers, pharmaceutical companies, government agencies to retail stores – are looking for experts to help them deal with data. The growing interest and demand for data

scientists is largely driven by the convergence of two factors. First, an unprecedented amount of data is being collected, from internet browsing logs, mobile telephony, medical imaging, genetic sequencing, and even personal information such as the number of steps taken or heart rates over time. The second factor driving this sector is the emergence of new computationally efficient and easy-to-use software tools that can extract useful patterns from this vast morass of raw data. These tools are the result of fundamental developments in statistics, computer science and machine learning which, when coupled with technological advances in the underlying hardware, lead to new computational techniques capable of analyzing extremely large and unusually complex data sets.

Almost every scientific research project, business operation, and government action generates, collects and mines large amounts of data. Success very much depends on a ready supply of technologically literate employees able to perform sophisticated data analysis. As demand for new and better technology grows, demand for data scientists will grow as well. Also, as data structure becomes more complex and data volume becomes increasingly large, requirements for data science positions and skills have become more demanding. According to the Bureau of Labor Statistics, an MS degree is typically required to enter this field. To meet the enormous demand for data scientists in academia, industry, business and government, higher educational institutions have been racing to establish data science and related programs to train qualified graduates. Among the Big Ten Academic Alliance (BTAA) institutions, eight have BS and/or MS programs in data science and six have business-oriented MA or MBA programs in business analytics.

The proposed MSDS program will be the only MSDS program in the state of Iowa. It will help the University of Iowa remain competitive nationally in the broad field of data science education and research and offer more data science training opportunities to the students in Iowa and the nation.

Relationship to existing programs at the institution. The proposed MSDS program does not duplicate existing programs at the University of Iowa. The most closely related existing master's degree is the Master of Science in Business Analytics (MSBA) program offered by the Tippie College of Business. It is a full-time business analytics program, offering courses in data visualization, databases, programming languages such as R and python from the business analytics perspective, and application areas such as text analytics, financial analytics, and healthcare analytics. This program trains students in advanced analytics techniques. It emphasizes developing skills to make students adept at using predictive and prescriptive analytics to solve business problems. This program is sixteen months and includes an internship, capstone project, and an emphasis on developing skills for clearly communicating results.

The university's Department of Computer Science (CS) offers a MS in Computer Science. In particular, the MS in CS has a machine learning and artificial intelligence bent to data science from the computer science perspective.

Finally, the Interdisciplinary Graduate Program in Informatics (IGPI) has the aligned goals of providing students the background and skills to work/study at the interface of computing and specific discipline (for instance, bioinformatics, information science, geoinformatics, health informatics, and library science). The focus of IGPI is application oriented, emphasizing data acquisition, data manipulation, database management and networking.

The proposed MSDS program is distinct from the aforementioned programs in the following ways:

- (a) In contrast to domain-specific analytics programs, for instance programs in business analytics, health analytics, and informatics, the proposed MSDS program is a generalist degree designed for students with a liberal arts BS degree and a certain

mathematical/analytical background (i.e., have taken and done well in two college calculus courses, one linear algebra course, one introductory computing course and have some basic coding skills in a computer programming language, such as C, Python and R). This target applicant pool is different from the applicant pool of MSBA, whose admission does not require college calculus and linear algebra;

(b) The emphasis and the course design of the proposed MSDS program is to train students on statistical thinking and modeling aspects of data science, so that they not only learn the existing data science methods and technologies, but are also capable of learning and developing new technologies after graduation; and

(c) The employment opportunities of the graduates from the proposed MSDS program are positions in academe, industry or government that involve experimental design, data visualization, and wrangling, modeling and analysis.

The proposed MSDS and the CS, MSBA, and IGPI programs constitute complementary programs for training the entire spectrum of personnel needed for fully harnessing the data revolution to the benefits of the state of Iowa and the nation. Moreover, the education of the students from these diverse programs will be greatly enriched as the programs collaborate by allowing students to take elective courses across programs.

The proposed MSDS program may also be beneficial to students in several other related programs such as biostatistics, educational measurement and statistics, and geographical and sustainability sciences at the University of Iowa. The students in these programs are often involved in data-intensive work in their training and research, and may enroll in some of the courses in the MSDS program.

Relationship to existing programs at other colleges and universities. There are no existing graduate MSDS programs at other colleges and universities in Iowa. There are two undergraduate programs in data science: The College of Liberal Arts and Sciences at Iowa State University offers a BS degree in data science. The University of Northern Iowa has a BA in physics with a data science emphasis program in the Department of Physics. The proposed MSDS is the only program with an emphasis on statistical modeling and analysis in the state.

Iowa State University's Ivy College of Business offers the Master of Business Analytics program, an online program for employed professionals that addresses the challenges of dealing with data analytics and business intelligence. It differs from the proposed MSDS program, which is a residential, generalist degree designed for students with a liberal arts BS degree.

Resources to establish a high-quality program. The personnel, facilities, and equipment currently in use for students within our existing graduate programs in the Department of Statistics and Actuarial Science, will also support the proposed MSDS, with ability to serve additional students at no extra cost. Specialized facilities are not required for the major.

Student demand. There is strong demand for data science training from undergraduate and graduate students across many subject fields at the University of Iowa and other universities in the nation and abroad because of the excellent employment and career opportunities.

For example, there has been a generally sustained increase in the enrollment of students from many disciplines in two existing data science courses, namely, Statistical Learning (STAT: 4540) and Data Visualization and Data Technologies (STAT: 4580). Over 65% of the students enrolled in these courses were from disciplines other than statistics or data science. The sustained interest of students in these courses, both from within and outside of the discipline, suggests that offering the proposed MSDS program will help meet student demand for data science training. The data

below indicating that data science is a growing field in the workforce in Iowa and beyond also indicate that the proposed MSDS program will be attractive to many students from Iowa, as well as the nation and the world.

Workforce need/demand. Data science has been one of the fastest growing industries over the past few years and, with the ever-increasing demand for data, the field is poised to grow for years to come. Indeed, data science has only just begun to expand as a discipline into industries and government agencies other than technology. For example, in 2019, the United States Office of Personnel Management for the first time recognized the job title “Data Scientist” as an official role in government agencies. In June the same year, the Department of Education National Center for Education System released the new Classification of Instructional Programs (CIP) codes for data science, data analytics and applied statistics.

According to the Bureau of Labor and Statistics employment growth of computer information and research scientists, which include data scientists, from 2019 to 2029 is 15%. PricewaterhouseCoopers (PwC) describes data science as among the most sought-after positions in America in their report on investing in America’s data science and analytics talent. Data science jobs ranked No. 2 on Glassdoor’s list of best jobs in America 2021 with median base salary of \$113,736. Data scientists have also been ranked as No. 3 in LinkedIn’s Emerging Jobs Report 2020 with an estimated 37% annual growth rate. Data from job boards like LinkedIn, Glassdoor and Dice.com consistently show strong job demand for data scientists. A search for “data scientists” (performed on February 13, 2021) on LinkedIn, Glassdoor, and dice.com yielded 27,568, 29,354 and 27,545 data scientist jobs, respectively. Such demand exceeds the number of data scientists that can be trained in all the existing data science and related programs in higher education institutions. Finally, there is also strong interest in data science from undergraduate and graduate students across many subject fields because of the excellent employment and career opportunities.

Funding and Cost. No special new faculty, facilities, and equipment are required for this proposed program since the faculty, required classrooms, and office space are currently adequate for this program.

The new program largely relies on existing courses. New resources specifically intended to support this program are not needed at this time. The new capstone course required for the program will be developed and offered by the existing faculty of the Department of Statistics and Actuarial Science.

	Total Costs
Year 1	\$0

Projected student enrollment.

Graduate	Yr 1	Yr 2	Yr 3	Yr 4	Yr 5	Yr 6	Yr 7
Majors	20	40	40	45	50	50	55
Non-Majors	5	5	5	10	10	10	10

In the early years of the program, we anticipate that most of our students will come from the undergraduate population at the University of Iowa, especially students from data science, computer science, mathematics, statistics and actuarial science, as well as students with a BS degree in any subject area who have taken and done well in calculus and linear algebra courses, and have some background in computer programming. Because the MSDS promises to be an attractive option for students pursuing bachelor’s degrees in certain areas (e.g., mathematics,

computer science, actuarial science, statistics, physics, chemistry and biology) and who may be interested in obtaining both the BS and MS degree in data science in 5 years, we will also pursue approval for an eventual undergraduate-to-graduate option.

After the program is established, we also expect to attract students from across the nation and abroad, since data science is in high demand nationally and in many countries of the world.

Accreditation. The program will not seek accreditation since such accreditation is not offered solely in the field of data science.

Evaluation plan. To assess the level of interest and sustainability of the program, the department will assess annually the number of applicants, enrollments, and retentions against the targets of the degree program; with the target applicants being 50% more than enrollments and the target retention rate greater than 80%. Program leadership will also track placement of our graduates (initial and five years out) to ensure our graduates are competitive in the workforce, finding jobs, and employed in occupations that they are enthusiastic about and find satisfying.

The College of Liberal Arts and Sciences will review the proposed MSDS at the third and fifth year to assess whether it attains its target enrollments and projected completion rates. The department chair will meet with the Associate Dean for Graduate Education and the Graduate Educational Policy and Curriculum Committee (GEPCC) of the College of Liberal Arts and Sciences to discuss these matters, with the program's faculty advisory committee then making recommendations on changes, if indicated.

The College of Liberal Arts and Sciences reviews each of its departments every seven years. At that time, both continuing and new programs of study offered by the department are reviewed as well. Particular attention is placed on size of the programs and curricular obstacles, if any, such as waitlists in key courses, courses with a low enrollment, and the number of students graduating with related degrees.

Date of implementation. Fall 2022.

Letters of Support

IOWA STATE UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Department of Statistics
1121H Snedecor Hall
2438 Osborn Drive
Ames, Iowa 50011-1090
515 294-7754
dnett@iastate.edu
dnett.github.io

August 15, 2021

Kung-Sik Chan
Robert V. Hogg Professor and Chair
Department of Statistics and Actuarial Science
The University of Iowa
241 Schaeffer Hall
Iowa City, Iowa 52242-1409

Dear Professor Chan:

As chair of the Department of Statistics at Iowa State University, I am writing to express my support for the University of Iowa's proposed Master of Science degree in Data Science (MSDS). The proposed MSDS program offers crucial training needed to extract information from the wealth of data in today's world. The program will help address scientific, industrial, and societal demands for personnel with expertise in data science. Students completing the proposed MSDS program should have excellent career opportunities in Iowa and throughout the nation.

Faculty in the Department of Statistics and Actuarial Science have the knowledge and training needed to deliver the proposed MSDS curriculum successfully. Furthermore, the proposed program appears to fit well with other programs offered at the University of Iowa, including the Bachelor of Science in Data Science, the PhD in Statistics with a data science track, and the Master of Science in Business Analytics. The proposed program also differs from existing programs at Iowa State University and, in particular, those currently offered by the Iowa State University Department of Statistics.

Thank you for giving me the opportunity to review the proposed MSDS program. I look forward to future communications and collaborations with your department in the areas of statistics and data science.

Sincerely,



Dan Nettleton, Ph.D.
Laurence H. Baker Endowed Chair
Distinguished Professor
Chair, Department of Statistics
Director, Baker Center for Bioinformatics and Biological Statistics



Department of Mathematics

August 10, 2021

Professor Kung-Sik Chan
Robert V Hogg Professor and Chair
Department of Statistics and Actuarial Science
The University of Iowa - 263 Schaeffer Hall
20 East Washington Street
Iowa City, IA 52242

Dear Kung-Sik:

Thank you for sharing the proposal for establishing a Master's program in Data Science (MSDS) at your department.

The Data Science industry is now well past the days when the ability to run a dozen or so machine learning algorithms with the help of R/Python packages was sufficient to enter the profession. Data Science is now much more advanced, complex, and demanding for a BS degree to be a satisfactory credential. A Master's degree in Data Science is now becoming more or less essential for good positions in data science. Your proposal to establish a MSDS program is, therefore, most welcome. It is very timely and your MSDS would be a great service to the data science industry.

Despite the inherently inter-disciplinary nature of data science, the core of data science has its largest overlap with Statistics. The relationship between the core and emerging data science methodologies and statistics is, in fact, becoming deeper and more fruitful day by day. I am very pleased that your department will administer the proposed MSDS program. You have a highly distinguished faculty with a world-wide reputation for cutting-edge research and education in computer-intensive 21st century Statistics. Your department's credentials are second to none. The MSDS curriculum details you have given fully reflect the expertise of your faculty, as well as the needs of data science professionals.

I am pleased to offer my strong support for your proposal and I wish you every success.

Regards,

A handwritten signature in black ink that reads "Syed Kirmani". The signature is written in a cursive style with a large, prominent "S" at the beginning.

Syed N.U.A. Kirmani
Professor of Mathematics

cc: Jose Herrera, Provost University of Northern Iowa
John Fritch, Dean College of Humanities, Arts and Sciences
Doug Mupasiri, Department Head
