1. **The name of the proposed degree/major**: MS Data Science
2. **Home department/department-like unit and School/College**: Department of Statistics, College of Letters and Science
3. **Mode of delivery** (face-to-face, online, other-specify): face-to-face
4. **Primary faculty contact person**: Dr. Yazhen Wang
5. **A description of the new program that includes anticipated program learning outcomes**

   The Department of Statistics, in conjunction with the Department of Computer Sciences, is proposing a new 131 program MS Data Science program. The program will be a joint effort between both departments with Statistics serving as the home department. This program will build upon expertise and existing programs within the new School of Computing, Data and Information Sciences (CDIS) to offer a Data Science graduate-level program that includes core courses from Statistics and Computer Sciences and electives from these and other departments to give graduates the ability to address advanced level data science challenges.

   Anticipate Program Learning Outcomes

   1. Demonstrates understanding of theories, methodologies, and computation as tools to solve complex problems in data science.
   2. Selects or adapts appropriate data science approaches and uses or develops best practices in data-driven applications.
   3. Synthesizes information, organizes insights, and evaluates impact pertaining to questions for studies involving empirical data.
   4. Communicates data science concepts and results clearly.
   5. Adheres to principles of ethical and professional conduct in data science.

6. **An overview of program content and program level and information about how the program may relate to other UW-Madison offerings.**

   The overall outline of the program intends to be a combination of content offered by the Department of Statistics and by the Department of Computer Sciences in a 30-credit MS degree.

   A curriculum is envisioned to consist of approximately ten courses, with 3 courses from Statistics, 3 from Computer Sciences, 2 courses in Machine Learning (to be selected from a roster of courses offered by both departments), and 2 electives. The electives could be additional courses in either Statistics or Computer Sciences, or suitable courses offered by other departments. The eight courses in Statistics, Computer Sciences, and Machine Learning will each be in a specified subject area, with several options for specific courses to be offered in each subject area.

The curriculum outlined above was arrived at by reviewing existing, successful programs and demands from employers. It is anticipated that some existing courses from both departments will be utilized, some existing courses will be adapted, and some new courses may be created.

The planning committee anticipates that prerequisites will include two semesters of calculus, a semester of linear algebra, and previous R or Python coursework. Additionally, they plan to explore opportunities for students to address deficiencies in linear algebra and Python through summer onboarding coursework.

Both the Statistics and Computer Sciences Departments offer professional, 131 MS programs, and they see this new MDS program as complementary to these existing programs. The MS Computer Sciences: Professional Master's Program is more general and spans a broad range of computer sciences topics. The MS Statistics: Data Science program is a data science oriented program but is focused on coursework in statistics and requires a strong quantitative background for admission. This new program would allow students to have a third option that provides training from both fields as it specifically relates to data science.

There are other existing programs on campus, through a variety of departments, that offer data science and data analytics graduate education. These programs are all "domain" based and are intended to teach students about data analytics or science within a specific context such as biomedicine, business, psychology, or engineering. For this reason the new MS Data Science program will not overlap with these programs.

7. **Existing or anticipated resources required to deliver the program; provide a summary of program support either from reallocation or from new resources, such as program revenue.**

   The Department of Statistics will be the administrative home and will collaborate with the Department of Computer Sciences to oversee the MDS program once it is created. The resources to deliver the program would be new resources based on the tuition paid by students. A revenue generating budget worksheet will be prepared with the full program proposal.

8. If relevant, **information on other required approvals to offer the program beyond the Board of Regents.** NA

9. **Explanation of how the new program aligns with the institutional mission, strategic plan, and existing academic degree program array**

   At the UW-Madison institutional level the proposed new program aligns with the new 2020-2025 strategic priorities. Specific priorities such as excellence in research and scholarship: ensure the continued vitality, competitiveness, and strength of their graduate and professional programs, excellence in teaching and educational achievement: expand educational programming in areas of high student demand, while maintaining the broad-based strength of the two departments' educational enterprise,

and a vibrant campus community: enhance diversity among their students, faculty, and staff and build upon their strong commitment to diversity to create a welcome, empowered, and inclusive community are paramount in the plans for the new program.

Within CDIS the new program will strive to meet the missions of contributing to the public good and enriching civil society through researching, teaching and outreach; educating responsible leaders, critical thinkers, and creative innovators; and fostering prosperity in the state, region, and the world.

These priorities are expected to be achieved through offering a highly ranked program based on best practices within data science, where new technologies and tools are implemented within the curriculum. Additionally, there is a plan to offer a summer "onboarding" option to students who are excellent candidates, but perhaps lack an important prerequisite course or two to increase the diversity of students in the program.

10. **A link to the institution's academic strategic plan**: https://strategicframework.wisc.edu/

11. **A rationale that clearly defines the need for the new program, including market and student demand in the context of local, regional, and UW System wide programs.**

    Data science is one of the fastest growing employment sectors in the nation and in Wisconsin.

    Although there is not a specific category labeled "Data Scientist" in the Occupational Outlook Handbook (OOH) from the Bureau of Labor Statistics, the job outlook in the period 2019-29 for Statisticians is projected to grow by 35% and for Computer and Information Research Scientists to grow by 15% (both classified as "much faster than average"). Both areas list the "typical entry-level education" for these fields to be a Master's degree. It is reasonable to predict that students graduating from a program such as this would be well positioned to meet continued demand for employees with these skills.

    The Division of Continuing Studies has provided us with an Occupation Analysis from Burning Glass Technologies for Data Scientists in the midwest (information retrieved December 2020). In this report one can see within the job metrics the projected growth is +19.5% for the next 10 years, average demand is 2,619 postings during 2020, and the median salary is $111,304. In analyzing technical skills necessary to meet this demand, topics such as data science, python, machine learning, R, and predictive modeling are already among the most desired skills and have projected growth for demand in the next 2 years. The proposed program will provide education to students in these important areas.

    Looking at the UW System major mania page there is currently only one other MS Data Science program available. This appears to be a collaborative online MS program across Eau Claire, Green Bay, La Crosse, Oshkosh, Stevens Point, and Superior. The proposed plan is to offer a face-to-face program that will build upon the existing

expertise at UW-Madison within CDIS, so there will be ample demand to allow for a financially viable 131 program that does not compete with other programs.

12. **A list of the program faculty who are central to the planning process and who will participate in the program with it is implemented.**

    - Remzi Arpaci-Dusseau - Computer Sciences, Department Chair
    - Shivaram Venkataraman - Computer Sciences
    - Yazhen Wang - Statistics
    - Stephen Wright - Computer Sciences
    - Brian Yandell - Statistics
    - Jun Zhu - Statistics, Department Chair

13. **Letters of support or concurrence from departments, schools/colleges, and other units that have an overlapping or substantial interest.**

    - Statistics
    - Computer Sciences
    - iSchool
    - Mathematics
    - Biostatistics and Medical Informatics
    - Business Analytics
    - Electrical and Computer Engineering
    - Industrial and Systems Engineering
    - Psychology