DTS 23262

UNIVERSITY OF HAWAI'I
BOARD OF REGENTS

'23 APR 26 P 1 :45

College of Natural and Health Sciences
Office of the Dean

# UNIVERSITY of HAWAI'I® HILO

April 3, 2023

## MEMORANDUM

TO:     Randolph G. Moore
        Chair, Board of Regents

        Ernest Wilson
        Chair, BOR Academic and Student Affairs Committee

VIA:    David Lassner
        President

VIA:    Debora Halbert
        Vice President for Academic Strategy

VIA:    Bonnie Irwin
        Chancellor

        Bonnie D Irwin
        Digitally signed by Bonnie D Irwin
        Date: 2023.03.22 14:54:40 -10'00'

VIA:    Kris Roney
        Vice Chancellor for Academic Affairs

        Digitally signed by Kris Roney
        Date: 2023.03.21 14:49:14 -10'00'

FROM:   Brian Wissman
        Interim Dean, College of Natural and Health Sciences

        Digitally signed by Brian Wissman
        Date: 2023.03.17 13:30:03 -10'00'

SUBJECT: REQUEST FOR PROVISIONAL STATUS FOR THE BACHELOR OF SCIENCE IN DATA SCIENCE AT THE UNIVERSITY OF HAWAI'I AT HILO

## SPECIFIC ACTION REQUESTED

It is respectfully requested that the Board of Regents grant provisional status to the Bachelor of Science in Data Science (BS-DSCI) in the College of Natural and Health Sciences at the University of Hawai'i at Hilo (UH Hilo).

## RECOMMENDED EFFECTIVE TERM/YEAR:
Fall 2024

## ADDITIONAL COST:
An additional operating cost of $12,000/year for a data science budget to purchase program specific software, maintain servers and equipment, branding/advertisement material, and other costs associated with the program. As the data science program matures and expands offerings the program anticipates the need for a faculty hire in year four.

200 W Kāwili St
Hilo, Hawai'i 96720-4091
Telephone: (808) 932-7506
Fax: (808) 932-7295
hilo.hawaii.edu/academics/cnhs
An Equal Opportunity/Affirmative Action Institution

## PURPOSE:
The provisional establishment of a Bachelor of Science degree in Data Science is being requested to provide students a program that teaches them the skills needed to collect, process, store, analyze and visualize data to drive decisions. This proposed degree supports data science as a key focus of President Lassner's post-pandemic vision for the UH system.

## BACKGROUND:
This request is in accordance with Board of Regents policy 5.201 III(A)(1) which states that "Approval of the board is required for the establishment of all new instructional programs granting academic credit leading to a degree or credential...." The ATP was approved by President Lassner on April 20, 2020.

## Significance/Contribution of this degree:
The Bachelor of Science in Data Science (BS-DSCI) at the University of Hawai'i at Hilo is designed as an interdisciplinary major that teaches students the computational and statistical principles needed to work with data from various sources. This proposed BS-DSCI degree builds up the current 18 credit data science certificate into a major with four domains of emphasis: Astronomy, Business, Computational, and Statistics. The interdisciplinary nature of the plan lends the program to be both a successful standalone degree for students, both first time freshman and transfer, and an excellent option for a secondary major.

Enrollment in computer and data science programs throughout the country are high with demand likely to increase. Data science skills are in demand and developing this program is in line with Governor Ige's 'Hawai'i 2.0 Digital Economy' recovery plan, the 2020-25 Statewide Science and Technology Plan, and the State Comprehensive Economic Development Strategy. These documents show that Hawai'i's economy is in need of diversifying, an area such as data science is a promising avenue for this diversification as there are both local and remote opportunities available for Hawai'i residents with this skill set.

This BS-DSCI would be the first data science major within the UH system, which will draw students to both the UH System and UH Hilo. One of UH Hilo's strengths is utilizing the Island of Hawai'i as a "living laboratory" - analyzing data from researchers will continue this laboratory concept into the data science realm.

## Demand projections:
Following the current strong demand for the current certificate in Data Science at University of Hawai'i at Hilo, we project that the Bachelor of Science in Data Science will enroll around 30 students with eight graduating by year six.

## Similar programs at other UH campuses:
The UH system does not currently offer a Bachelor of Arts or Bachelor of Science in Data Science, but other campuses do offer degrees with data science concentrations.

UH West Oʻahu offers a Bachelor of Arts in Business Administration with a specialty in Data Analytics and UH Mānoa recently launched a data science track within their Computer Science major.  Currently, UH Hilo has a certificate in data science.  Outside of the UH System, Chaminade University offers a Data Science major.

Cost and resource allocation/reallocation implications:
Four current faculty at UH Hilo (Computer Science, Mathematics, Marine Science, and Business) were hired as part of the EPSCoR Data Science project.  As a condition of their hire, these faculty are obligated to teach at least one data science course and support the development of the degree program.  Leveraging these faculty along with two already allocated new hires in computer science, the data science program will be covered through the initial years.  As enrollment and demand for courses grow, additional faculty may need to be hired in the cognate areas (e.g. Mathematics) to cover those areas of need.

Impact of new program/program change request on campus budget allocations and mission priority:
The BS-DSCI program aligns with the 2021-2031 UH Hilo Strategic Plan.  It places emphasis on the investing in data to build a student-ready equity-minded campus culture, developing high-impact practices in more collaborative interdisciplinary opportunities that improve student success.

ACTION RECOMMENDED:
It is respectfully recommended that the Board of Regents grant provisional status to the Bachelor of Science in Data Science (BS-DSCI) in the College of Natural and Health Sciences at the University of Hawaiʻi at Hilo.


Attachments
   Provisional Program Proposal
   Letters of Support


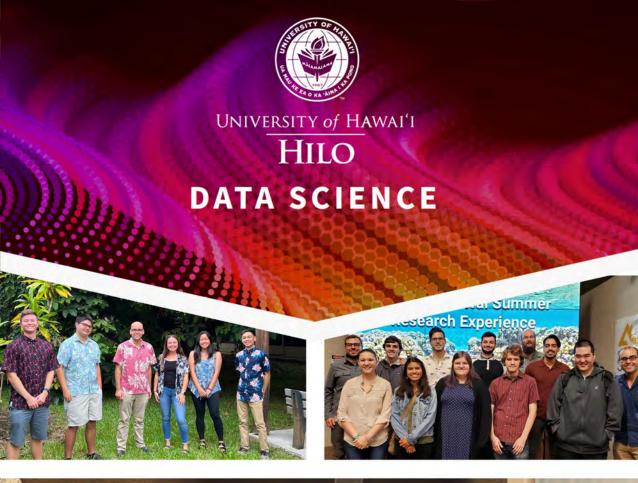c:      Interim Executive Administrator and Secretary of the Board Jamie Go

# UNIVERSITY of HAWAIʻI
## HILO
# DATA SCIENCE

## Table of Contents

## Executive Summary

The data science program at UH Hilo is designed to be an interdisciplinary major which teaches students the skills needed to collect, process, store, analyze, and visualize data to drive real-world decisions.  The core courses in our Data Science major teach students the computational and statistical principles needed to effectively work with data of various kinds and complexity (including images, text, tabular data, etc.). Our proposed major contains four tracks (Business, Astronomy, Computational, and Statistical) in which students receive more in-depth training needed to tackle data science problems in a specific domain.  We plan to add more tracks to the major over time as student interest and enrollment increases. At just under 50 credits, Data Science is designed to be attractive to a wide variety of students, particulary transfer students.

## Program Need and Demand

Data science skills are in high demand across the country and the state. This effort is in line with Governor Ige's 'Hawai'i 2.0 Digital Economy' recovery plan, the 2020-25 Statewide Science and Technology Plan (SSTP), and the State Comprehensive Economic Development Strategy (CEDS). These documents point out that Hawai'i's economy is in dire need of diversifying beyond the traditional reliance on tourism; data science is a very promising avenue as it allows workers to collaborate remotely with companies on other islands or the mainland while earning enough to live comfortably in Hawai'i. The Bureau of Labor Statistics projections for data scientists nationally show an anticipated 36.6% growth in the sector by 2031. As the SSTP observes, however, "data scientists" alone does not capture the scope of occupations that are touched by data science: "Employers that are reported to have engaged with emerging educational programs in DS [data science] during the development of this plan include (but are not limited to): finance, banking and insurance, defense contracting, public utilities, natural resource managers, non-profits, sports teams, small businesses, agriculture, energy, academic researchers across domains, acute and community healthcare facilities and systems, transportation providers, tourism, educators, social non-profits, travel and tourism, the criminal justice system and the military" (10). The breadth of where data analysis appears has only grown since the SSTP was completed in just 2020. The proposed BS in Data Science includes four subject matter tracks and a core data science curriculum that will help to train our students and build a workforce to address these market demands.

Hawai'i must be capable of training skilled data scientists to avoid the potential of a digital divide that leaves the state (government, employers, community) lagging both in data science jobs and the use of data to empower our growth and resilience. This urgency is illustrated by initiatives such as the Association of Hawaiian Civic Club's 2019 Resolution to support and advocate for data science that addresses Hawaiian advancement and integrates the Native Hawaiian community into these data endeavors.

Data science is also a key area of focus in President's Lassner's post-pandemic vision for the UH system; however, while Chaminade University has a dedicated Data Science major, the UH System does not. UH Manoa recently launched a Data Science track in their computer science major, and UH West Oahu has a Bachelor of Arts in Business Administration with a specialty in Data Analytics, and UH Hilo presently has only a certificate in Data Science.  UH Hilo has established itself as a "living laboratory" where ecologists, marine scientists, geologists, astronomers and more study many phenomena on the island - data science directly supports these research activities, which are increasingly requiring advanced data analytics. The BS proposed here capitalizes on these strengths in the inclusion of subject-matter application tracks. Moreover, UH Hilo's small and diverse departments make it easy to collaborate across fields to foster an interdisciplinary program like data science (which combines computer science skills, mathematics skills, and skills and knowledge of a particular domain).

As an emerging and interdisciplinary field of study, data science is also conducive for acquiring extramural funding to increase UH's research competitiveness and provide opportunities for students to have high-impact research experiences or internships. Funding bodies such as the National Science Foundation heavily support data science research and education – one of their '10 Big Ideas' is "Harnessing Data for 21st Century Science and Engineering." In fact, the four core data science faculty have successfully acquired grants totaling $ 4,166,410.00, as well as other collaborative awards (such as advocating for over $289K of the first year of the recent EPSCoR award to be allocated to UH Hilo to support data science research needs).   Much of this funding is specifically allocated to engaging UH Hilo undergraduates in data science research experiences. These experiences integrate research and education, helping students learn how their data science skills can directly address real-world research problems that affect Hawaiʻi and the world.   Currently, since data science is only a certificate, we have more data science research slots open than we have qualified students. This degree program is key to aligning our curriculum with the skills students will need to solve real-world problems and advance their careers in science and technology. Extramural support is not limited to just research activities: the Academy of Creative Media currently provides $60,000 in annual funding to UH Hilo to support data science education in next-generation digital visualization activities that provide students with instruction and training in cutting-edge data visualization tools. These ACM supported courses illustrate how UH Hilo is developing a valuable educational niche by offering an innovative data science curriculum that is relevant in the digital media workforce. The interdisciplinary nature of the Data Science degree program enables students to access resources (e.g., computer labs, clubs, equipment) from the wide variety of departments involved in the Data Science major, including both those departments that offer foundational courses as well as those that offer electives through the different domain emphasis tracks. This interdisciplinary exposure can be beneficial and we anticipate many students will decide to pursue a minor or dual major in a complimentary subject.

Additionally, the BS meets the imperatives outlined in the University of ʻʻi Strategic Plan, particularly those of "Meet Hawaiʻi's workforce needs for today and tomorrow" and "Diversify Hawaiʻi's economy through UH innovation and research." The means by which these imperatives are achieved overlap with the alignment with all three goals of the 2021-2031 UH Hilo Strategic Plan:

> **Goal 1: "UH Hilo will achieve a thriving student-ready equity-minded campus culture by investing in data."** Data-driven insights have the potential to improve student success, equity, and diversity. But this requires not just a single data analyst or office, but rather a culture of data analysis which is shared among the faculty, staff and students. This program will allow us to grow a robust and interdisciplinary data science pipeline, developing such a culture of data generation and analysis throughout the university.
> **Goal 2: "Making high-impact educational practices available to each student results in more collaborative interdisciplinary opportunities that improves student retention and overall student success."** High-impact practices are a centerpiece of our interdisciplinary data science program. Faculty in our program consistently offer data science research experiences for our students throughout the school year and summer, as well as external data science internships through programs such as Akamai.
> **Goal 3: "UH Hilo will support and invest in its campus ʻohana through … fostering interdisciplinary and collaborative efforts that achieve student success."** The data science program is designed to be highly interdisciplinary, allowing students to receive strong CS and math skills and then apply these skills to address challenging problems facing Hawaiʻi, whether they be in Astronomy, Business, Ecology, etc.

Finally, the Data Science program has broad faculty support. Of the 30 biennium budget proposals submitted to UH Hilo Faculty Congress in September 2022, the proposal to fund hires to support the Data Science program was ranked #3 by the voting members of faculty congress. The more highly ranked proposals (#1 and #2) were for computer science classroom equipment which indirectly benefits the Data Science program (as most of data science classes are taught in computer science classrooms). Additionally, the approved FY23 budget for UH Hilo includes funds for a new tenure-track assistant professor in computer science and a new instructor in computer science, advancing UH Hilo's ability to deploy both an expanded computer science curriculum and to support data science. Likewise, Marine Science and Business are presently recruiting new assistant professors, which will also increase capacities in both those programs and in data science.

## Enrollment and Completion Projections

Enrollment has been consistently strong in the existing Data Science certificate courses, and the certificate has been quite popular.  Since its inception in 2020, 25 students have completed the program. In fact, for UH Hilo students who graduated in Spring 2022, the Data Science certificate was the most popular certificate awarded by the College of Natural and Health Sciences.  It is notable that students seeking the certificate are enrolled in an array of majors, including not only Computer Science but also Marine Science, Environmental Science, Biology, Mathematics, and Astronomy.   In fact, During the Fall 2021 semester, despite its limited resources the CS department had to offer two sections of CS172 (one of the gateway data science courses) to meet demand from astronomy and computer science majors. In addition, the physics and astronomy department is in the process of changing both of its majors to require CS172, which will only increase demand further.  Additionally, our gateway classes (DATA 101 and DATA 200) will help interest students in the major in a gentle manner while satisfying GE requirements.

Based on estimates of conversions from the certificate, as well as new enrollments of primary and secondary majors, the anticipated enrollment with the average STEM-major year-to-year retention rate is as follows.

| Enrollment Projections:  Provisional Years | | | | | | |
|---|---|---|---|---|---|---|
| | **Year 1** | **Year 2** | **Year 3** | **Year 4** | **Year 5** | **Year 6** |
| | 10 | 17 | 28 | 32 | 35 | 43 |

Based on estimates of average completion within STEM majors (averaged over 5 years), the projected completion rate will reach the target of 10 or more graduates at year 6, just after the first program review.

| Program Completion Projection | | | | | | |
|---|---|---|---|---|---|---|
| | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 | Year 6 |
| Projected Completion | 0 | 1 | 2 | 3 | 6 | 15 |

## Resources

The existing facilities are sufficient to launch the Data Science program.  With the newly approved hires for computer science, we do not anticipate needing new faculty for data science in the initial years of the program. As enrollment increases there will likely be a need for a new tenure-track position in MATH, as there is currently only one tenure-track faculty qualified to teach the math-focused data science courses. Lastly, we would request a dedicated data science budget ($12K/year) to assist with tasks such as

purchasing data science software, maintaining data science servers and equipment, generating appropriate branding and advertising material for the program, etc.

| Existing Resources/Funding | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 | Year 6 | Provisional Total |
| Combined Tuition/ Summer/ Course Fees | 73,440 | 124,848 | 209,664 | 244,224 | 267,120 | 328,176 | 1,247,472 |
| Other Allocation (Grants) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| New Personnel Costs | | | | | | | |
|---|---|---|---|---|---|---|---|
| Personnel (incl. 5% increase annual) | Year 1 FY 24 | Year 2 FY25 | Year 3 FY26 | Year 4 FY27 | Year 5 FY28 | Year 6 FY29 | Total New |
| Projected New Tenure-line Faculty FTE | 0 | 0 | 0 | 1.0 | 0 | 0 | 1.0 |
| Projected New Lecturers FTE | 0 | 0 | .25 | .25 | .25 | .25 | .25 |
| Projected Reallocations FTE | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Projected New Salaries (w/fringe for lecturers) | 0 | 0 | 26,612 | 107,950 | 29,340 | 61,614 | 22,5516 |
| Projected Student Support (tutors) | 5,000 | 5,000 | 5,000 | 5,000 | 5,000 | 5,000 | 30,000 |
| Cumulative New Salaries | 5,000 | 5,000 | 31,612 | 112,950 | 118,340 | 154,814 | 402,716 |

| New Operating Costs | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 | Year 6 | Total New |
| New Operating Costs (new equipment and refresh) | 12,000 | 12,000 | 12,000 | 12,000 | 12,000 | 12000 | 72,000 |

The four EPSCoR Data Science faculty located in CS, Math, Marine Science, and Business will constitute the primary faculty assigned to the new Data Science major, as each is already obligated to teach at least one course in data science and support the development and success of the program. However, the CS and Math departments will share most of the burden and will need to devote more faculty resources over time to teach data science courses as opposed to the traditional courses in their discipline. For Computer Science, this need will be met through the 2.0 FTE hires approved and funded by the legislature in 2022. These positions will either contribute to teaching the data science courses directly, or reallocating the load for existing faculty. For Math, in the short-term lecturers will be used to reallocate faculty workload toward data science the

needed faculty resources, with permanent hires in place once demand for the relevant math courses has been established.

## Consultation between campuses

The system-wide effort to establish data science education programs aligns with having a full major and operational program with sufficient infrastructure at UH Hilo.  There has been extensive consultation between campuses as part of the previous and current UH System-wide EPSCoR award.  The last two awards have specifically focused on data science, and one of the focuses of the last two awards has been building a data science major at UH Hilo.  This has involved substantial presentation of our current certificate program and future major plans to the involved campuses in the UH System.  Specifically, data science faculty at UH Manoa (Jason Leigh, Peter Sadowski, Madhi Belcaid) are members of the current EPSCoR award and thus aware of the coming proposal; Indeed, they have noted that the UH Hilo Data Science major will feed nicely into the planned Data Science graduate program at UH Manoa. We have notified multiple programs across the UH System which teach curriculum relating to data science education to ensure there are no conflicts and to solicit support for developing the Data Science major at UH Hilo (see letters of support in Appendix).

## Associated Risks

- **Faculty recruitment and retention:** Data science expertise, especially in areas such as machine learning, is in high demand in the workforce, with many industrial research labs and data science startups heavily recruiting top PhDs in the field.  As such, it can be hard to recruit high-quality candidates who can greatly contribute to data science teaching, research, and service - these individuals are in short supply and often prefer to work in industry or at large research universities. Recruiting efforts will be widely advertised and carefully crafted to emphasize the strengths of UH Hilo, including teaching excellence, a diverse and productive academic community, the opportunity to shape a small program, and exciting opportunities to collaborate with scientists studying natural phenomena unique to our Island.  Despite recruitment efforts, failed searches (or searches which result in CS faculty without expertise in data science) are a possibility, as are faculty departures. Budget planning for future hiring will include likely high demand differentials estimated differentials. Further, we will mitigate the risk of faculty departures through a strong mentoring network in which the tenured data science faculty will provide advice on successful teaching approaches and shared teaching resources, opportunities for research collaboration, grant-writing advice and support, and support in crafting a successful dossier for contract renewal and tenure.

- **Program Marketing:** Given that data science is a relatively new field, students may not be aware of the program when searching for colleges and selecting an initial major. We propose a threefold strategy to combat this:

    o **Continuing students:**
        ▪ Gateway courses to attract students who might otherwise be unaware of the discipline or feel ill-prepared to data science while fulfilling GE requirements
        ▪ Broad data science inclusion in general education and preparatory curriculum: 100-level courses (CS172) required for other majors (such as Physics and Astronomy) will highlight the data science major and draw in interested students to both the major and the certificate.

    o **Freshmen enrollment:** Prioritize advertising the data science program through admissions and events. Advertisements should highlight the uniqueness in Hawaiʻi, the job opportunities, the interdisciplinary nature, the faculty expertise, and the associated research opportunities. The current team, for instance, has NSF-funded data science projects related to Artificial Intelligence, Computer Vision, Marine Science, Climate Science, and Text Mining all of which are actively hiring paid undergraduate research assistants over the school year and during the summer.

    o **Transfer students:** Advisement, events, and pathways for students with UHCCs to assure awareness of the data science major and its relatively low number of required credits for the major, and great job opportunities. Faculty will partner with admissions to develop targeted advertising for transfer students and seek to maximize enrollment from this group of students.

- **Long-term Physical Facility Needs:** The B.S. in Data Science will not require new facilities, as current Math/CS facilities are suitable for housing the Data Science program. However, as the program grows its enrollment there will be a need for more classroom and engagement space that can be repurposed from existing UH Hilo space. Although there are two classrooms where data science classes are prioritized (College Hall 5 and LRC 350), no space is currently dedicated to data science research and undergraduate student activities. Investing in dedicated facilities for this program will better position UH Hilo for success by allowing us to attract world-class faculty and students by showcasing our institutional support for data science. As the campus refreshes the campus master plan, these areas of note will be included in the needs assessment.

## Program Structure and Effectiveness

### Staffing

The courses will be taught by the existing data science faculty members as well as by new hires in computer science. The new CS faculty members, 1 tenure-track professor and 1 instructor, provide the additional expertise and time that the data and computer science majors need to grow capacities and continue to broadly support students. The core data science faculty will take responsibility for program assessments and curriculum.

| Learning Outcomes | |
|---|---|
| Short Title | |
| Math | Understand the mathematical foundations of data science, including probabilistic reasoning as well as Bayesian and Frequentist statistics. |
| Programming | Independently create computer programs which analyze complicated real-world datasets, as well as modify data science programs written by others. |
| AI/ML | Apply machine learning and artificial intelligence to effectively solve real-world problems. |
| Tools | Identify the optimal data science tools required for various analytical procedures and data visualization tasks. |
| Limits | Understand the appropriate usage and limits of data science, e.g. what kinds of questions can be asked and answered versus those that cannot be addressed. |
| Types | Manipulate common data types: e.g., traditional experiments, image collections, natural language text, and real-time time-series processing. |
| Security | Recognize privacy and security issues surrounding data of various types. |
| Formats | Retrieve and store data in different formats, using databases as and popular file formats. |
| Domain | Draw on knowledge from numerous disciplines, including those outside of mathematics and computer science such as business, natural sciences, health, and social sciences. |
| Software | Use existing data science and informatics software effectively. |
| Communication | Communicate insights from large datasets to others, in written, oral, and visual forms. |

### Assessment

Program effectiveness will be assessed through multiple means:

1. Programmatic assessment: The data science team has crafted a comprehensive plan for WASC-compliant programmatic assessment of student learning outcomes in the data science program (see above). Details available in Appendix: Assessment Matrix and Assessment Plan.
   a. We have mapped out 11 student learning outcomes for the major to the 17 core data science courses as part of the major (excluding the tracks), indicating High, Medium, and Low Coverage of each

SLO. This allows us to quickly identify which courses should be teaching students each of our learning outcomes

    b. For each of the 11 learning outcomes, we have created an assessment plan. Six of the learning outcomes will be evaluated by selecting exam questions from the relevant courses, which a group of data science faculty will score using rubrics. Two of the learning outcomes will be evaluated by scoring student's programming and data science projects using a targeted rubric. The remaining three of the learning outcomes will be evaluated by scoring a student's written reports, and/or oral presentations at the end of the culminating capstone course.

    c. If student performance is unexpectedly low for a certain learning outcome, the data science group will seriously discuss this and quickly develop strategies to improve student performance, whether that be through revising the way a certain subject is taught, or even adjusting our major to require additional courses that teach a certain topic.

2. External Review - Periodically, the data science faculty will invite data science faculty at other institutions to visit UH Hilo to review the program. These external faculty will visit classes, hold meetings with faculty and students, and generate a written report outlining the strengths of the program and giving constructive criticism. Evaluations will be included as a portion of annual assessments and cumulative program reviews.

3. Alumni Surveys: The degree coordinator will collect permanent emails from graduating students (if this email address is later found to be inactive, faculty will use LinkedIn profiles, etc. to contact the students). After graduation from the Data Science Program, students will receive emails to determine what they have lined up next, to quantify the percent of students that acquire employment in the field of study (both in Hawai'i and elsewhere). We will, in partnership with the Institutional Research Office, create a short survey to gather their responses to 1) employment, 2) employment in field, 3) location of employment (Hawai'i or elsewhere), 4) perceptions of their experience in the program, and 5) perceptions of the program's value to them in the workplace. The survey will go out every three years.

4. Research measures: Number of research proposals submitted under the umbrella of data science, number of students involved in undergraduate data science research (during the academic year and through summer research experiences)

5. Enrollment data and graduation outcomes, including persistence and time to degree.

- 7 math foundations credits (calculus 1 + matrix)
- 12 lower-division credit
- 20 upper-division credits
- 8-9 credits focused in specific discipline

Data Science Recommended GE courses that serve as potential entryways (optional for degree):
- DATA 101 Awesome data science skills (3)
- DATA/QBA 200 Intro to Business Analytics (3)

Math Foundations: (7 credits)
- MATH 241 Calculus 1 (4)
- MATH 211 Matrix Algebra With Applications (3)

Core Requirements: (23 Credits)
- DATA/CS 171 Data Science Fundamentals in R (3)
- DATA/CS 172 Python for Data Analysis (3)
- DATA/CS 272 Machine Learning for Data Sci (3)
- DATA/MATH 271 Applied Statistics with R (3)
- DATA 315 Mathematical Methods for Data Sciences (3)
- CS 475 Data Visualization (3)
- CS 475L Data Visualization Lab (1)
- DATA 495 Data Science Seminar (1)
- DATA 490 Data Science Capstone (3)

Data Management and Analysis: Choose one course from the below (3 Credits)
- DATA/CS 373 Data Security & Privacy (3)
- DATA 474 Applied Informatics (3)
- DATA/CS 370 Data Management (3)
- DATA/MATH 371 Multivariate Modeling with R (3)

Types of Data: Choose two courses from the below (6 Credits)
- DATA/CS 483 Computer Vision (3)
- DATA/QBA 465 Text Mining for Behavior and Social Sciences (3)
- DATA 470 3D Mapping of Ecosystems (3)

Domain Emphases (8-9 Credits)
Pick one of the following "domain emphases:

Computational Track (9 Credits)
- CS 350 Systems Programming (3)
- CS 440 Artificial Intelligence (3)

- CS 422 Database Analytics (3)

Business Track: (9 Credits)

Choose 9 credits from:
- QBA 260 Business Statistics (3)
- QBA 300 Operations Management (3)
- QBA 362 Business Analytics (3)
- QBA 367 Business Analytics and Applications (3)

Statistics Track (9 Credits)

Choose 9 credits from:
- MATH 311 Linear Algebra (3)
- MATH 371 Multivariate Stats (3) (See notes)
- MATH 421 Probability (3)
- MATH 422 Mathematical Statistics (3)

Astronomy Track (8 Credits)

Choose 8 credits from:
- ASTR/PHYS 260 Computational Physics & Astron (3)
- ASTR/PHYS 260L Computational Phys & Astr Lab (1)
- ASTR 250 Observational Astronomy (3)
- ASTR 250L Observational Astronomy Lab (1)
- ASTR 385 Software Systems for Astronomy (3)

Program Notes:
1. A C or better is required in all courses required for the major
2. At least 45 credits must be earned in courses at the 300- and 400- level.
3. Math 371 not eligible for the statistics track if already taken for the data management and analysis course.

| Anticipated Courses, Sections, SSH | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 | Year 6 | Current Year |
| No. New Courses Offered | 0 | 4 | 4 | 4 | 0 | 1 | 0 |
| No. New Sections Offered | 1 | 4 | 4 | 4 | 0 | 1 | 0 |
| Annual SSH | 240 | 300 | 400 | 600 | 700 | 800 | 0 |

## Appendix

4 Year Map for BS in Data Science
Assessment Matrix
Assessment Plan
Extramural Funding
Letters of Support

| 4 Year Map for B.S. in Data Science | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Year 1** | | **Year 2** | | **Year 3** | | **Year 4** | |
| **Fall** | | **Fall** | | **Fall** | | **Fall** | |
| MATH 241 (FQ: QUANT REASONING) | 4 | DB: BIO SCI | 3 | DA, DH or DL: ARTS, HUM, LIT II | 3 | GE GCC | 3 |
| GL: LANG ARTS | 3 | DS: SOC SCI I | 3 | GE HPP | 3 | WI Elective | 3 |
| DATA/CS 171 | 3 | "Types of Data" class 1 | 3 | DATA 315 | 3 | DATA 496 | 1 |
| DATA/CS 172 | 3 | DATA/MATH 211 | 3 | Domain Emphasis class 1 | 3 | Domain Emphasis class 3 | 3 |
| Elective | 3 | Elective | 3 | Upper Elective | 3 | Upper Elective | 3 |
| | | | | | | Upper Elective | 3 |
| Credits | # | Credits | # | Credits | # | Credits | 16 |
| **Spring** | | **Spring** | | **Spring** | | **Spring** | |
| ENG 100 (FW: WRIT COM) | 4 | DP: PHYS SCI | 3 | WI Elective | 3 | Upper WI Elective | 3 |
| DATA/MATH 271 (GQR: QUANT REASONING) | 3 | DY: SCI LAB | 1 | DS: SOC SCI II | 3 | DATA 490 | 3 |
| FGA, FGB or FGC: GLOBAL MULTICUL II | 3 | DA, DH or DL: ARTS, HUM, LIT I | 3 | CS 475 | 3 | FGA, FGB or FGC: GLOBAL MULTICUL I | 3 |
| DATA/CS 272 | 3 | Data Management class | 3 | CS 475L | 1 | Upper Elective | 3 |
| | | "Types of Data" class 2 | 3 | Domain Emphasis class 2 | 3 | Upper Elective | 3 |
| | | Elective | 3 | Upper Elective | 3 | | |
| Credits | # | Credits | # | Credits | # | Credits | 15 |
| Credits | 0 | Credits | 0 | Credits | 0 | Credits | 0 |
| | | | | | | | 12 |
| Total Credits | # | Total Credits | # | Total Credits | # | Total Credits | 2 |

**Upper Division Credits: 47**

**FAQ:**

If you say:

**FQ: QUANT REASONING**

- Then any FQ course can count

**MATH 241 (FQ: QUANT REAS)**
- Then MATH 241 is required, and will also count for FQ

**FQ: QUANT REASONING (Rec: MATH 125)**

- Then any FQ course can count

- MATH 125 is Recommended
- If a course is Recommended, a student is not required to take it

**Program Notes:**

| | |
|---|---|
| 1 | "Types of Data" - students must take at least 6 credits from DATA/QBA 465, DATA/CS 483, and DATA 470 |
| 2 | "Data Management" - students must take at least 3 credits from CS 373, MARE 474, CS370, MATH 471 |
| 3 | Domain Emphasis - students must take at least 8-9 credits from a single domain track.  Students may choose between four tracks: Computational, Business, Statistics, and  Astronomy |
| 4 | A C or better is required in all courses required for the major |
| 5 | At least 45 credits must be earned in courses at the 300- and 400- level. |

| Assessment Matrix | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Math | Programming | AI/ML | Tools | Limits | Types | Security | Formats | Domain | Software | Communication |
| Course title: | | | | | | | | | | | |
| MATH 241: Calculus 1 | H | | | | | | | | | | |
| MATH 211 Matrix Algebra With Applications | H | | | | L | L | | | | | |
| DATA/CS 171 Data Science Fundamentals in R | L | L | | H | | L | L | M | L | M | M |
| DATA/CS 172 Python for Data Analysis | | H | | L | | L | L | H | L | | |
| DATA/CS 272 Machine Learning for Data Sci | M | M | H | L | H | L | M | L | L | | M |
| DATA/MATH 271 Applied Statistics with R | H | L | | M | H | L | | L | L | M | M |
| DATA 315 Mathematical Methods for Data Sciences | H | M | M | L | H | L | | | | | |
| CS 475(L) Data Visualization (Lab) | | L | L | M | | L | | L | M | M | H |
| DATA 495 Data Science Seminar | H | | H | | | | | | | | H |
| DATA 490 Data Science Capstone | M | H | M | M | M | M | M | M | M | M | H |
| DATA/CS 373 Data Security & Privacy | M | M | L | M | | | H | | | | |
| DATA 474 Applied Informatics | | M | L | H | | | | M | M | H | L |
| DATA/CS 370 Data Management | | M | | | | | | | | H | |
| DATA/MATH 371 Multivariate Modeling with R | H | L | | M | M | M | | | M | L | L |
| DATA/CS 483 Computer Vision | M | M | H | M | | H | | M | L | | |
| DATA/QBA 465 Text Mining for Behavior and Social Sciences | | L | H | M | | H | | M | M | H | |
| DATA 470 3D Mapping of Ecosystems | | L | L | H | | H | | M | M | H | |

*Notes:* highest reinforcement occurs in domain emphasis tracks

## Assessment Plan

| Learning Outcome | Assessment Tool or Measure | Target or Benchmark | Timing | Who is Responsible for Assessment and Analysis? | Information Flow for Use and Dissemination | [Consider: Analysis & Results, Actions, Follow-up] |
|---|---|---|---|---|---|---|
| Math | Exam Questions | Students can solve basic statistical problems without assistance | Every 4 years, starting year 1 | Grady Weyenberg | Grady collects responses for a class like MATH 241 and creates rubrics, John, Travis, Grady, Sukhwa grade questions, results disseminated to all DS faculty | Trends over time will be noted. A dip in performance will lead to exploration of more effective ways to teach our math/statistics classes, with much focus being on MATH241 |
| Programming | Exam Questions | Students can reason about and improve code in Python without assistance | Every 4 years, starting year 1 | Travis Mandel | Travis collects responses for a class like CS 172 and creates rubrics, Travis, Grady, Sukhwa grade questions, results disseminated to all DS faculty | Trends over time will be noted. A dip in performance will lead to exploration of more effective ways to teach our intro programming classes, with much focus being on CS 172 |
| AI/ML | Exam Questions | Students can determine what ML framework is appropriate for a variety of problems | Every 4 years, starting year 2 | Travis Mandel | Travis collects responses for a class like CS 272 and creates rubrics, Travis, Grady, Sukhwa grade questions, results disseminated to all DS faculty | Trends over time will be noted. A dip in performance will lead to exploration of more effective ways to teach our AI/ML classes, with much focus being on CS272 |
| Tools | Written reports that students will create at the end detailing the methodology and results behind a semester-long data science project (Capstone Course) | Students can effectively utilize existing tools as part of a larger capstone project (and effectively describe the use of tools in a written report) | Every 4 years, starting year 4 | John Burns | John creates rubrics for students final reports in capstone class (focusing on how tools are integrated into the projects), John, Travis, Grady, Sukhwa evaluate reports according to rubrics | Trends over time will be noted. A dip in performance will lead to exploration of more effective ways to teach utilization of data science tools, with focus classes being CS171, and DATA474 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Limits** | Exam Questions | Students can effectively articulate the limits of supervised machine learning algorithms | Every 4 years, starting year 3 | Travis Mandel | Travis collects responses for a class like CS 272 and creates rubrics, Travis, Grady, Sukhwa grade questions, results disseminated to all DS faculty | Trends over time will be noted. A dip in performance will lead to exploration of more effective ways to teach limits of data science, with much focus being on CS 272 |
| **Types** | Exam Questions | Students can compare contrast/ ways to analyze natural language text data from other types of data (i.e. tabular) | Every 3 years | Sukhwa Hong | Sukhwa collects responses for a course like QBA 465 and creates rubrics, Travis, Grady, Sukhwa grade questions, results disseminated to all DS faculty | Trends over time will be noted. A dip in performance will lead to exploration of more effective ways to teach about various types of data, with focus being on courses like CS483 and QBA 465 |
| **Security** | Exam Questions | Students can identify data security, privacy, and ethics concerns relating to data collected from human subjects | Every 3 years | Sukhwa Hong | Sukhwa collects responses for a course like QBA 465 or CS373 and creates rubrics, Travis, Grady, John,Sukhwa grade questions, results disseminated to all DS faculty | If performance is insufficient, we may consider requiring CS373 to better emphasize this key learning outcomes relating to privacy and security |
| **Formats** | Programming assignment involving using Python to analyze datasets stored as databases, CSVs, and/or JSON files | Students can correctly utilize Databases, CSV files, and/or JSON files | Every 3 years | Travis Mandel | Travis collects student submissions for a course like CS 172 and creates rubrics, Travis, Grady, Sukhwa grade questions, results disseminated to all DS faculty | If performance is insufficient, we may consider requiring CS370 to better emphasize this key learning outcomes relating to databases and data storage |

| | | | | | | |
|---|---|---|---|---|---|---|
| Domain | Written reports that students will create at the end detailing the methodology and results behind a semester-long data science project (Capstone Course) | Students synthesize domain-specific knowledge to solve frame data science problems and describe steps taken in a written report | Every 3 years | Sukhwa Hong | Sukhwa creates rubrics for students final reports in capstone class (focusing on how domain science is presented and how it informs the projects) , Travis, Grady, John, Sukhwa grade reports from capstone course according to rubrics, disseminated to all DS faculty | If performance is insufficient, will break the data down by domain emphasis track to decide which tracks need improvement.  We will also consider the match between the capstone project and the track the student studied. |
| Software | Digital project in which visualization software is applied to a data science problem in a domain like marine science, ecology, etc.  (Such projects are a common part of courses like CS475) | Using existing data visualization and analysis software effectively | Every 3 years | John Burns | John creates rubrics for a visualization-related course project in a class like MARE 474 or CS 475, focusing on how they effectively use visualization software) Travis, Grady, John, Sukhwa grade report from capstone course according to rubrics, disseminated to all DS faculty | Trends over time will be noted. A dip in performance will lead to exploration of more effective ways to teach about various types of data science software, with  focus being on courses like CS 475 or MARE 474 |
| Communication | Final presentations in the Capstone Course in which students present their final data, science projects and results | Effectively present motivation, methodology, results, and conclusions | Every 3 years | Grady Weyenberg | Grady creates rubrics for presentations focusing on presentation skills and organization rather than technical content, Travis, Grady, John, Sukhwa watch presentations and rate according to reports, results disseminated to all DS faculty | Trends over time will be noted. A dip in performance will lead to exploration of more effective ways to teach about presentation skills, perhaps requiring more presentations earlier in the program to give students practice. |

As new faculty are hired, they will be integrated into assessment responsibilities.

| | Extramural Grants Acquired by Data Science Faculty | | |
|---|---|---|---|
| PI | Project Title | Funding Agency | Total Award Amount |
| Burns | Understanding environmental stressors and deep reefs to support management of Papahanaumokuakea | NFWF | $ 900,000.00 |
| Burns | E Hui Pu: A collaborative approach to understand climate change impacts on traditional Hawaiian Opelu (mackerel scad) aggregation sites in South Kona, Hawaiʻi | PICASC | $ 35,000.00 |
| Burns | National Science Foundation Supplemental Award to Understanding Biotic Response to Environmental Change in Tropical Ecosystems Through a Place-Based Context | NSF | $ 200,000.00 |
| Burns | Co-Knowledge Production of Historic Property and Habitat Assessments on Navy Coastal Submerged Lands for the Island of Guam | NAVFAC | $ 900,000.00 |
| Burns | Mitigating impacts from environmental stressor: early response to the Chondria outbreak at Pearl and Hermes Atoll, and Hurricane Walaka impacts at French Frigate Shoals. | NFWF | $ 300,000.00 |
| Burns | He ala aukai - The Path Near the Sea; Adapting to Climate Inflictions Upon Intertidal Shoreline. | PICASC | $ 40,000.00 |
| Burns | Burroughs Wellcome Fund: Postdoctoral Enrichment Program for Underrepresented Minorities | Burroughs | $ 120,000.00 |
| Burns | Academy of Create Media Strategic Investment Competition | ACM | $ 84,000.00 |
| Burns | Evaluate benthic monitoring protocol in Pacific Island and South Florida Caribbean inventory and monitoring networks. National Park Service | NPS | $ 287,000.00 |
| Burns | Benthic, Fish Biota, Anchialine Pond, and Intertidal Pool Surveys | NELHA | $ 32,000.00 |
| Burns | Coral response to land-to-ocean freshwater flux: A ridge-to-reef perspective. | USGS | $ 51,000.00 |
| Burns | Quantifying Vulnerability to Sea Level Rise Across Multiple Coastal Typologies | NASA | $ 525,000.00 |
| Burns | Honolulu Channel Monitoring Proposal | DAR | $ 44,000.00 |
| Burns | Capacity-Building for Division of Aquatic Resources to Integrate 3D Mapping Techniques into Monitoring of Hawaiian Nearshore Coral Reef Ecosystems | DAR | $ 90,000.00 |
| Mandel | CAREER: Accelerating Scientific Data Collection through Human-in-the-Loop Artificial Intelligence | NSF | $ 105,669.00 |
| Mandel | CAREER: Accelerating Scientific Data Collection through Human-in-the-Loop Artificial Intelligence | NSF | $ 104,876.00 |
| Mandel | CAREER: Accelerating Scientific Data Collection through Human-in-the-Loop Artificial Intelligence | NSF | $ 113,941.00 |

| | | | |
|---|---|---|---|
| Mandel | CAREER: Accelerating Scientific Data Collection through Human-in-the-Loop Artificial Intelligence | NSF | $ 115,117.00 |
| Mandel (co-PI), Weyenberg (Key Personnel) | SCC-PG: Big Island Drink Smart | NSF | $ 149,961.00 |
| Weyenberg | Collaborative Research: Principal Component Analysis over Treespaces and its Applications to Phylogenomics | NSF | $ 118,807.00 |
| | | | **$ 4,316,371.00** |
| **Budget From recent EPSCoR award allocated to UHH - Year 1** | | | |
| Burns, Mandel, Hong | Research Budget | NSF | $ 264,722.48 |
| Mandel | Participant Support Budget | NSF | $ 24,510.00 |
| | | | **$ 289,232.48** |