Utah System of Higher Education
New Academic Program Proposal
Cover/Signature Page - Full Template

Institution Submitting Request:	**University of Utah**
Proposed Program Title:	**Bachelor of Science, Data Science**
Sponsoring School, College, or Division:	**College of Engineering**
Sponsoring Academic Department(s) or Unit(s):	**School of Computing**
Classification of Instructional Program Code[1] :	**30.0801 : Mathematics and Computer Science**
Min/Max Credit Hours Required to Earn Degree:	**122 / 122**
Proposed Beginning Term[2]:	**Fall 2019**
Institutional Board of Trustees' Approval Date:	

Program Type (check all that apply):

☐ (AAS)	Associate of Applied Science Degree
☐ (AA)	Associate of Arts Degree
☐ (AS)	Associate of Science Degree
☐	Specialized Associate Degree (specify award type[3]:)
☐	Other (specify award type[3]:)
☐ (BA)	Bachelor of Arts Degree
☒ (BS)	Bachelor of Science Degree
☐	Professional Bachelor Degree (specify award type[3]:)
☐	Other (specify award type[3]:)
☐ (MA)	Master of Arts Degree
☐ (MS)	Master of Science Degree
☐	Professional Master Degree (specify award type[3]:)
☐	Other (specify award type[3]:)
☐	Doctoral Degree (specify award type[3]:)
☐	K-12 School Personnel Program
☐	Out of Service Area Delivery Program

Chief Academic Officer (or Designee) Signature:
I, the Chief Academic Officer or Designee, certify that all required institutional approvals have been obtained prior to submitting this request to the Office of the Commissioner.

Please type your first and last name _____ Date:

☐ I understand that checking this box constitutes my legal signature.

[1] For CIP code classifications, please see http://nces.ed.gov/ipeds/cipcode/Default.aspx?y=55
[2] "Proposed Beginning Term" refers to first term after Regent approval that students may declare this program.
[3] Please indicate award such as APE, BFA, MBA, MEd, EdD, JD

Utah System of Higher Education
Program Description - Full Template

Section I: The Request

University of Utah requests approval to offer the following Baccalaureate degree(s): Bachelor of Science, Data Science effective Fall 2019. This program was approved by the institional Board of Trustees on .

Section II: Program Proposal

Program Description

Present a complete, formal program description.

We plan to create one of the first comprehensive Data Science bachelor's degrees in the Mountain West. This program will prepare individuals to develop and apply knowledge of basic computer science and software engineering sufficient to be able to build, modify, or use software tools for data analysis. These students will also learn fundamentals of data analysis and processing in order to be able to effectively, efficiently, and ethically make decisions based on the information within various data sources.

Graduates of the proposed program will fill a growing demand for data scientists and work in a variety of industries including health care, finance, and the internet. They will typically work with consumers of data in order to analyze, manage, and augment large data sets or work in industries that require automated forms of decision making and analysis. In situations that require large and sophisticated software development for data analysis, these graduates may work with or lead teams of computer scientists or software developers in a joint effort.

Consistency with Institutional Mission

Explain how the program is consistent with the institution's Regents-approved mission, roles, and goals. Institutional mission and roles may be found at highereduutah.org/policies/policyr312/ .

The students obtaining the new degree would be positioned to be leaders in the emerging area of data science, helping grow the state of Utah's economy in this growing area. It would also organize and energize the transfer of knowledge from the University's faculty to its students, in this area where significant new methodologies are being actively developed.

Section III: Needs Assessment

Program Rationale

Describe the institutional procedures used to arrive at a decision to offer the program. Briefly indicate why such a program should be initiated. State how the institution and the USHE benefit by offering the proposed program.

The proposal of this new program is the result of a several year exploratory effort.
First, we have supervised a few Bachelor's of University Studies degrees that focus on data science, and like the proposed degree are centered on courses in the basic computer science (mainly from School of Computing), basic statistics (mainly from the Department of Mathematics), and advanced analytics (from School of Computing). They also were completed with courses from other quantitative domains with data specific focuses (e.g., business, economics). The students felt the degrees fulfilled the skills needed for their career paths, and have found jobs in their desired areas after graduation.

Second, the advanced analytics courses, which help differentiate this degree's requirements from those of a traditional bachelor's in computer science, have had significant increase in enrollment by undergraduate students in the last few years. These courses (e.g., Machine Learning, Data Mining, Visualization for Data Science) were originally designed as graduate level courses. Moreover, in an effort to allow undergraduates to take these courses, we have introduced a new class (now called "Foundations of Data Analysis") which can be taken by undergraduates at the sophomore level, and allow undergraduates to succeed in the previously graduate level courses now taken at the sophomore or junior level.

undergraduate enrollment:
Data Mining (CS 5140/6140) : S14 11 | S15 31 | S16 22 | S17 40
Machine Learning (CS 5350/6350) : F13 26 | F14 23 | F15 12 | F16 35 | F17 54
Visualization (CS 5630/6630) : F13 21 | F14 18 | F15 12 | F16 9 | F17 11
Database Systems (CS 5530) : S14 95 | S15 76 | S16 135 | S17 132

Third, on August 29, 2017, we hosted a focus group of 6 local data scientists, and presented a draft of this proposed new degree. The high-level summary is that
(1) "the degree would prepare students for future careers in data science" and
(2) the SoC "is recommended to pursue creating this degree."

Fourth, on November 17, 2017, the School of Computing had a faculty meeting to discuss the proposed degree. In an effort to get a sense of the faculty's opinion on the new degree, the faculty was unanimously in favor of pursuing the creation of these new degrees.

Fifth, the faculty of the College of Engineering Council met on October 19, 2018 and voted unanimously to approve the proposal.

Labor Market Demand

Provide local, state, and/or national labor market data that speak to the need for this program. Occupational demand, wage, and number of annual openings information may be found at sources such as Utah DWS Occupation Information Data Viewer (jobs.utah.gov/jsp/wi/utalmis/gotoOccinfo.do) and the Occupation Outlook Handbook (www.bls.gov/oco).

In recent years, Data Scientist is commonly listed as one of best jobs in the United States. For instance, a Forbes article from January 2018 (https://www.forbes.com/sites/louiscolumbus/2018/01/29/data-scientist-is-the-best-job-in-america-according-glassdoors-2018-rankings/) based on the hiring website "Glassdoor.com" reports
"Data Scientist has been named the best job in America for three years running, with a median base salary of $110,000 and 4,524 job openings." Related jobs "Analytics Manager," "Data Engineer," and "Data Analyst" rank 18, 33 and 38, respectively, on this list. Note that all of these jobs are listed as distinct from more traditional Computer Science jobs "Software Engineer," "DevOps Engineer," "Mobile Developer," "Front End Engineer," etc.

The Utah Economic Data Viewer does not include an explicit category for "data scientist," with the closest being "Computer and Information Research Scientists." This is likely related also to a Computer Science degree, but likely includes data scientists. For this field, the annual median state-wide wages are $91,090 and there are expected 30 annual openings statewide and 600 in the United states. While the listing notes that many of these jobs expect a "Doctoral or professional degree," the data scientist panel we hosted in August 2017 provided some insight into why an applicant with a BS in Data Science may be competitive. Many entry level jobs are labeled as "Data Engineer" and "Data Analyst" whereas jobs titled "Data Scientist" often expect some seniority and include some management component. At least among those on the panel, they mainly obtained this necessary experience on the job.

Student Demand

Provide evidence of student interest and demand that supports potential program enrollment. Use Appendix D to project five years' enrollments and graduates. Note: If the proposed program is an expansion of an existing program, present several years enrollment trends by headcount and/or by student credit hours that justify expansion.

We project a strong student demand (perhaps 10 or more will immediately transfer from other degrees). This is based on several factors.

* SoC Associate Professor Jeff Phillips has supervised 3 students (one is ongoing) in Bachelors of Undergraduate Studies degrees with emphasis in Data Science.
* The enrollment in the "Advanced Data Analysis" courses (listed above) which distinguishes the proposed Data Science degree from the Computer Science degree have had large *undergraduate* enrollments. They are mainly co-listed with graduate courses, but have had at least 9 undergraduate students. For Data Mining, Machine Learning, and Database Systems, this number has been typically at least in the mid 20s and growing; each had at least 40 undergraduates enrolled in 2017.
* In Fall 2017 the SoC helped launch a Data Science Club (in the process of registering with ASUU). This club's events occur roughly monthly, and has had a typically attendance of about 30-40 students. The opening event featuring a panel of data scientists attempting to explain this developing field and how to become part of it, had about 100 students. This club seems to be a roughly even mix of graduate (mainly MS) and undergraduate students.

Similar Programs

Are similar programs offered elsewhere in the USHE, the state, or Intermountain Region? If yes, identify the existing program(s) and cite justifications for why the Regents should approve another program of this type. How does the proposed program differ from or compliment similar program(s)?

Within the University of Utah, the most similar programs are

* BS in Computer Science:

This degree overlaps non-trivially with the proposed degree. They have the same (or very similar) pre-major requirements, and teach the same basic programming skills required for efficient data processing. The Computer Science degree, compared to the proposed degree, requires several additional courses in software engineering and low-level programming focusing on interactions with the hardware of the computer. These additional courses CS 3505 "Software Practice II", CS 3805 "Computer Organization", and CS 4400 "Computer

Systems" are some of the most demanding courses in the Computer Science degree. However, the CS degree requires only one course in linear algebra or probability/statistics, whereas the proposed degree will require 3-4 foundational courses across these areas. The CS degree in place required a foundational course in automata or numeric computation, not required and less relevant to the proposed degree.
The proposed degree also requires 5 additional courses: 2 in data management (CS 5630 "Database Systems" and a new course "Data Wrangling") and 3 in data analysis (CS 5140 "Data Mining", CS 5350 "Machine Learning", and CS 5630 "Visualization for Data Science").

* BS in Mathematics - Statistics Emphasis:
  This degree overlaps with the proposed degree in that it teaches much of the required mathematical foundations, and provides a strong basis in probability and statistics, it does not require any programming/computational training necessary for efficiently managing and processing data, and does not require the advanced data analysis techniques.

* BS in Information Systems:
  This degree covers the basics of data processing and management as well as statistics. However, since this degree also requires completing the Eccel's School "core requirements" which includes topics on accounting, finance, marketing, and business operations, it does not cover as much depth in the computational or statistical side. Moreover it does not require the advanced data analysis techniques.

Other Universities in the USHE offer some similarly related degrees.
* Utah State University offers degrees for instance in Computer Science and in Statistics. The USU Computer Science degree is no more similar to the proposed degree than the University of Utah one. The USU Statistics degree trains in some of the data analysis skills, but requires none of the computational ones.

We are aware of only two Data Science undergraduate degree in the Mountain West.

  * at BYU Idaho: http://www.byui.edu/catalog/#/programs/41PwqJ9RZ
  * at Montana Tech: https://www.mtech.edu/academics/clsps/data-science/degree.htm

These are similar to the proposed degree in that they blend computer science, statistics, and advanced data analytics.

There are numerous undergraduate degrees offered in the Mountain West in Computer Science, and Statistics, and Business Analytics at universities like BYU, Boise State, and UC Boulder, but as with the ones described above, they are not data science degrees.

For instance, BYU offers an Applied & Computational Mathematics Emphasis within their Math degree; this still requires very little training in programming and advanced data analysis techniques like Machine Learning, Data Mining, Databases, and Visualization.

However, across the country, several universities have been creating new degrees in Data Science:

  * The College of Charleston has offered a BS in Data Science for over a decade:
    http://datascience.cofc.edu/program-information/
  * The Ohio State University offers a Data Analytics Major, jointly administered by the Computer Science and Statistics Departments:
    https://data-analytics.osu.edu/major/core-curriculum
  * University of California-Irvine offers a BS in Data Science, administered within the Statistics Department, but with many CS courses:
    http://www.stat.uci.edu/slider/b-s-in-data-science/
  * The University of Rochester offers a BA and BS in Data Science, run through their Goergen Institute for Data Science:
    http://www.sas.rochester.edu/dsc/undergraduate/major.html
  * Penn State University offers degrees in Data Sciences with Applied, Computational, and Statistical Modeling emphasis; it is an inter-college initiative.
    https://datasciences.psu.edu
  * UC-Berkeley is preparing to offer a major in Data Science, but with many of the course plans already in place; including an introductory course with 1700 students!
    https://data.berkeley.edu/education/faqs

These degrees are all similar to our proposed degree, in that

they blend computer science training, with statistical training, and with new advanced techniques (e.g., Machine Learning, Data Mining, Visualization). What department runs the degree typically depends on which offers these advanced courses; they are all in the School of Computing at the University of Utah.
More similar degrees can be found here: (https://www.bachelorsportal.com/study-options/269779226/data-science-big-data-united-states.html). The ones listed above are those from the more highly-ranked universities.

Finally, it should be noted that many universities offer one or more graduate-level degrees that are in Data Science. The naming conventions are fairly broad (at the University of Utah, the most similar degree is the "MS in Computing - Data Management and Analysis Track"). The proposed degree differentiates from all of these in that it offers the same sorts of training at the undergraduate level (albeit sometimes, as compared with the SoC MS degree, covering a bit less advanced material).

Collaboration with and Impact on Other USHE Institutions

Indicate if the program will be delivered outside of designated service area; provide justification. Service areas are defined in higheredutah.org/policies/policyr315/ . Assess the impact the new program will have on other USHE institutions. Describe any discussions with other institutions pertaining to this program. Include any collaborative efforts that may have been proposed.

We plan to only offer this degree at the University of Utah, and have no plan to collaborate with other USHE Institutions.

External Review and Accreditation

Indicate whether external consultants or, for a career and technical education program, program advisory committee were involved in the development of the proposed program. List the members of the external consultants or advisory committee and briefly describe their activities. If the program will seek special professional accreditation, project anticipated costs and a date for accreditation review.

There are no current accreditation standards for Data Science degrees. However, certain guidelines have been drafted (including some with aid of involved faculty); we plan to watch as these guidelines and our proposed program matures, and may seek accreditation if one is standardized in the future.

We will continue to stay in contact with local data scientists through other efforts in this area, and the SoC's Industrial Advisory Board. We expect to continue to seek feedback on the

degree, including hosting additional focus groups in the future as the degree matures.

Section IV: Program Details

Graduation Standards and Number of Credits

Provide graduation standards. Provide justification if number of credit or clock hours exceeds credit limit for this program type described in R401-3.11, which can be found at highereutah.org/policies/R401.

The degree will require between 102 and 109 credit hours specific to the major.
12-15 in PreMajor Requirements
27 in General Ed Requirements (including one class specific to the degree)
17-18 in Analytical Foundations
10 in Computing Foundations
12 in Core Data Science (Advanced Data Science Techniques)
9 in Elective Advanced Data Analysis
9-12 in Elective Data Domain
6 in Capstone Project
Students will be required to earn at least 122 credit hours total to complete the Bachelor of Science degree.

A C- or better will be required in all PreMajor, Analytical Foundations, and Computing Foundation courses. A 2.5 GPA average will be required among the Core Data Science courses and among the Elective Advanced Data Analysis courses. Students can retake a course at most twice for it to be counted towards these grade point requirements.

Admission Requirements

List admission requirements specific to the proposed program.

This major is eligible for any student admitted as an undergraduate at the University of Utah. The pre-major courses are open to all students, and must be completed to apply for full-major status.

We plan to set the requirements for full major status at the same level as the BS in Computer Science. The pre-major courses mirror those in Computer Science, and initially will be the exact same courses. Moreover, numerous of the required courses are CS courses which require full major or minor (in Computer Science) status to be enrolled (unless permission is given by the instructor). Thus if we deviate from these standards, it would allow a backdoor into the advanced Computer Science courses. These standards will also ensure the students

are fully prepared for the advanced courses which have been otherwise taught at the graduate level and may be on the challenging side.

The requirements for applying to full major status in Computer Science are detailed here:
http://www.cs.utah.edu/how-to-apply/
and are roughly to complete pre-major courses with minimum C-, and average 3.0 or higher in those courses and overall at the University. A committee then evaluates all qualified candidates, and fixes a standard at which to grant or deny full-major status. The standard is set so the number of students granted full-major status is set based on a controlled, but aggressive growth rate, while also having confidence the students will succeed in the more advanced courses required to complete the degree. The Data Science bachelors will plan to set a similar status for full major status in Data Science by conferring with the Computer Science committee.

Curriculum and Degree Map

Use the tables in Appendix A to provide a list of courses and Appendix B to provide a program Degree Map, also referred to as a graduation plan.

Section V: Institution, Faculty, and Staff Support

Institutional Readiness

How do existing administrative structures support the proposed program? Identify new organizational structures that may be needed to deliver the program. Will the proposed program impact the delivery of undergraduate and/or lower-division education? If yes, how?

All of the required/elective courses are already taught regularly by faculty in the School of Computing or Mathematics, with one exception. The courses which are not required in existing programs are popular electives, so they are generally taught anyways. Two of these courses:

CS 3190: Foundations of Data Analysis
CS 6964: Ethics of Data Science

were designed with this major in mind, and to meet other student and academic demands.

The one new course planned would be a course on Data Wrangling, taught for students at roughly the sophomore level. Even if not part of this new degree program, there seems to be demand for this course as an elective for computer science undergraduates. A course taught in health sciences, overlaps

fairly significantly with the planned Data Wrangling course:
  MDCRC 6521: Intro to Computer Programming
and cross-listed as
  BMI 6018: Intro to Programming.
We are discussions to coordinate between this existing course and the planned one, with the potential to cross-list them.

If demand for certain courses increases, it may dictate that some courses are taught every semester instead of once a year.

The School of Computing and Department of Mathematics are in discussions regarding aligning content between
  CS 3130 / ECE 3530 : Probability and Statistics for Engineers
and
  Math 3070 : Applied Probability I
These courses may be cross-listed, among other options. Either course will fulfill the same requirements for the proposed degree - the program lists Math 3070 for simplicity.

The initially will plan is to have all BS in Data Science students to complete a senior project or senior thesis in a two course, 6 credit hour sequence. The plan is to have them take the same courses CS 4000 / CS 4500 for a project or CS 4940 / CS 4970 for a thesis as the Computer Science major students. The difference is that the project will require a more data science focused theme. As the demand grows for this major, distinct versions of these courses will be created for the Data Science students.

Faculty

Describe faculty development activities that will support this program. Will existing faculty/instructions, including teaching/ graduate assistants, be sufficient to instruct the program or will additional faculty be recruited? If needed, provide plans and resources to secure qualified faculty. Use Appendix C to provide detail on faculty profiles and new hires.

The existing faculty in the School of Computing and Department of Mathematics will be sufficient to teach these required and elective courses, and overall support the program. All of the required and elective courses are already taught in some form on a regular basis, except one planned course (Data Wrangling). The SoC is committed to regularly teaching such a class on a regular basis starting soon. Moreover, the TEP in Statistics (between SoC, Mathematics, and BioEngineering) will already plan to hire 1 additional faculty in addition to the recent SoC hire (started Spring 2018) and 1 recent Math hire (started Fall 2019) in this area.

Staff

Describe the staff development activities that will support this program. Will existing staff such as administrative, secretarial/ clerical, laboratory aides, advisors, be sufficient to support the program or will additional staff need to be hired? Provide plans and resources to secure qualified staff, as needed.

The existing staff at the School of Computing and Department of Mathematics will be sufficient to support this program. In fact, the SoC recently hired an additional undergraduate advisor in Spring 2018.

Student Advisement

Describe how students in the proposed program will be advised.

Students accepted into the Data Science major will be assigned an undergraduate advisor within the School of Computing, and will be advised through similar methods as has been used successfully within the SoC. Pre-Major students will have access to the same advising staff as Pre-Major students with interest in Computing or Mathematics, but will now be advised on details of this major option.

In particular, the advisors in SoC will incorporate this new major option into their advising. When they meet with new students, they go over the requirements needed to apply for the major and suggest which courses to take so they have a clear path and schedule. They provide a class availability sheet so students can plan around when courses are offered and what the prerequisites are, so they can plan accordingly. When freshmen students attend orientation they go over the program requirements for the CS CS EAE majors they offer, and will add the DS major requirements into this repertoire. For recruitment, they will include the DS major into events such as Engineering Day, Major Expo, and Red White and U Day.

A very helpful part of this advising has been the CS-major "green sheet" which provides all course requirements on one sheet, with possible 4-year paths through the major on the reverse side. We have created a draft version of such a sheet in planning the DS major and will have one to help in advising and recruiting students.

Library and Information Resources

Describe library resources required to offer the proposed program if any. List new library resources to be acquired.

No additional library resources will be needed on a recurring basis at this time. In this area, many of the text books are

available freely online. If new texts are adopted, individual faculty using those text books will work with the library to ensure they are available.

Projected Enrollment and Finance

Use Appendix D to provide projected enrollment and information on related operating expenses and funding sources.

Section VI: Program Evaluation

Program Assessment

Identify program goals. Describe the system of assessment to be used to evaluate and develop the program.

We will assess the degree along three main vectors.

First, SoC plans to track the number of students who are formally accepted into the major, and how many graduate. It is expected that this number to start at about 10 students per class, and hopefully it will continue to grow from there.

Second, SoC will track the rate at which the program's graduates get jobs, and which get jobs within the data science domain. Similar to students graduating with a BS in Computer Science, it is expected that virtually all students will get technical jobs upon graduation. There is less confidence in how accurately SoC can track the type of jobs students obtain. As learned from discussions with local data scientists, including from the August 2017 focus group, often entry level jobs in this area are titled something else, such as ``data analyst'' or ``data engineer,'' and the title ``data scientists'' is often reserved for someone more senior.

Despite that, through a new Data Science mailing list, since Fall 2017, the involved faculty have announced about a dozen openings for jobs with titles clearly in the data science domain, several titled "data scientist."

Third, SoC hopes to use the opportunity of launching a new STEM +C degree, to reset perceptions of the type of student who this degree is for. In particular, the hope is to draw from a more diverse audience than other STEM+C degrees. The story telling and tangible, human aspects of data science has the potential to attract a wider class of students. As a baseline, the hope is to compare favorably to other STEM+C degrees at the University of Utah in terms of traditionally underrepresented groups.

We will assess this by tracking the percentage of underrepresented

groups (e.g., females, minorities) in comparison to the CS and other STEM+C degrees.

Student Standards of Performance

List the standards, competencies, and marketable skills students will have achieved at the time of graduation. How and why were these standards and competencies chosen? Include formative and summative assessment measures to be used to determine student learning outcomes.

Upon graduation with a BS in Data Science, it is expected for students to be able to have the following skills.
(1) To efficiently manage, process, and compute with a wide array of data types. This will be achieved through both high- and low-level programming languages. The graduates should be able to interact with software engineers who build infrastructure around these tasks, as well as consumers and producers of the data. They should also be prepared to learn new technology, such as new programming languages or new software tools, as they becomes useful in their domain.
(2) To master probabilistic and statistical thinking. When faced with a new task or a decision, they should be able to recognize the non-determinism of the outcome, and assess with some confidence the likelihood of various potential outcomes.
(3) To take an abstract task involving a data set, and perform standard data analysis. These include, identifying the core structure and patterns of the data if it exists, building a function that automatically predicts the outcome for as of yet unseen data, and communicating concisely these findings to consumers of the data.
(4) To interact meaningfully with experts in some technical data domain. These experts may be for instance engineers, scientists, marketers, or policy setters. In particular, our graduates should have some non-trivial experience interacting with the finer details of at least one data rich domain.

Assessment of these desired learning outcomes will occur through a few forms. We will review and track course evaluations and curriculum of the relevant courses to make sure they continue to align with our goals. These roughly align as:
For skill (1) the Computing Foundations and Database courses.
For skill (2) the Statistics and Machine Learning courses.
For skill (3) the Data Analysis, Data Mining, and Visualization courses.
For skill (4) the elective Data Domain courses and the Capstone project where the students will be required to synthesize all of these skills into a large cohesive data science task.
In particular, the capstone project will culminate in a forum and

presentation of each project by students. Then the application and demonstration of each specific skills will be directly assessed by industrial and faculty judges at the Capstone presentation forum. They will be given a sheet asking them to judge the demonstration of each of these skills. We will aggregate these survey results to assess the overall health of the program.

We will also hold exit surveys and personal interviews with a small subset of graduating students. We plan to then again interview alumni 3-5 years after graduation after working in the field. Additionally, we will communicate with the professors teaching those data domain courses to gauge how the DS students are learning, interacting with the domain-focused students, and applying their data processing and analysis expertise in the contact of those courses. Finally, we will interact with the new employers of the DS graduates through the SoC IAB, the Utah Data Science Day, as judges at the Capstone presentation, and other similar events.

These four skills are the distillation of the 4 main components of the proposed degree. The first is derived from the overlap with the BS in Computer Science. The second is based on the overlap with the Statistics emphasis part of the BA in Mathematics. The third part are the advanced courses which overlap with the successful Graduate Certificate in Big Data, and are the modern techniques most associated with data science. The fourth will force students to go beyond merely focusing on the technical aspects, and become aware of some of the less well-defined and communication aspects of this profession.

This desired skill set was presented to the August focus group via a draft of the degree, and it was agreed upon that each was a core element of what would make a data scientists successful. These aspects are also found in the various emerging guidelines for data science programs. Two examples are

* “Curriculum Guidelines for Undergraduate Programs in Data Science” (https://arxiv.org/abs/1801.06814). This report was generated by the Park City Math Institute 2016 Summer Undergraduate Faculty Program which met for the purpose of computing guidelines for undergraduate programs in data science.
* “Data Science for Undergraduates: Opportunities and Options (2018)” (https://www.nap.edu/catalog/25104/data-science-for-undergraduates-opportunities-and-options). This monograph was

published by the National Academies of Science, Engineering, and Medicine, and reviewed by leading computer scientists, statisticians, applied mathematicians and others in data-driven scientific domains. Its goal is to provide “considerations and approaches for academic institutions and others in the broader data science communities to help guide the transformation to this field.“

The data science faculty will meet annually in early Fall to review the above assessment criteria and data, and to assess whether any nationally developed curriculum guidelines should be more formally adopted. We will also ensure at least one representative for data science is part of the SoC undergraduate curriculum committee so that the same standards and review is held for data science courses, and so the needs of the DS students and degree planning can be planned for.

Appendix A: Program Curriculum

List all courses, including new courses, to be offered in the proposed program by prefix, number, title, and credit hours (or credit equivalences). Indicate new courses with an X in the appropriate columns. The total number of credit hours should reflect the number of credits required to be awarded the degree.

For variable credits, please enter the minimum value in the table for credit hours. To explain variable credit in detail as well as any additional information, use the narrative box at the end of this appendix.

Course Number	NEW Course	Course Title	Credit Hours
General Education Courses (list specific courses if recommended for this program on Degree Map)			
		General Education Credit Hour Sub-Total	24
Required Courses			
CS 1030		Foundattions of Computer Science	3
CS 1410		Object-Oriented Programming	4
CS 2420		Algorithms and Data Structures	4
Math 1310		Engineering Calculus I	4
Math 1320		Engineering Calculus II	4
CS 2100		Discrete Structures	3
Math 2270		Linear Algebra	4
Math 3070		Applied Statistics I	4
Math 3080		Applied Statistics II	4
CS 3190		Foundations of Data Analysis	3
CS 4962		Ethics and Data Science	3
CS 3300	X	Data Wrangling	3
CS 3500		Software Practice I	4
CS 4150		Algorithms	3
CS 5140		Data Mining	3
CS 5350		Machine Learning	3
CS 5530		Database Systems	3
CS 5630		Visualization for Data Science	3
CS 4000		Senior Capstone Design	3
CS 4500		Senior Capstone Project	3
		Required Course Credit Hour Sub-Total	68
Elective Courses			
Math 5080		Statistical Inference I	3
Math 5090		Statistical Inference II	3
CS 4300		Artificial Intelligence	3
CS 5150		Advanced Algorithms	3
CS 5190		Probabilistic Learning	3
CS 5340		Natural Language Processing	3
CS 5530		Human Computer Interaction	3
CS 5640		Image Processing	3
CS 6530		Database Systems	3

Course Number	NEW Course	Course Title	Credit Hours
		Elective Credit Hour Sub-Total	18
		Core Curriculum Credit Hour Sub-Total	110

Program Curriculum Narrative

Describe any variable credits. You may also include additional curriculum information.

The curriculum for the proposed degree can be decomposed into 8 categories:

1. Premajor Requirements (12-15 credit hours):
This component provides introductory programming and mathematical foundations, and mirror the premajor requirements for a BS in Computer Science. Among the CS requirements CS 1030, CS 1410, and CS 2420, the first course is often tested out of for Computer Science majors, and the proposed degree would offer the same option. The Calculus I & II requirements (Math 1310 and 1320) or near equivalent (Math 1210 and 1220) are standard for most quantitative majors. As the Data Science degree matures, SoC may elect to offer a
DS 1030, Foundations of Data Science
as a optional replacement for CS 1030 that focuses more on data science applications at the very introductory levels while also acquainting students with computational thinking.

2. General Education Requirements (27 credit hours)
As in all undergraduate degrees, the proposed degree will require Wrtg 2010 (WR2), Wrtg 3012/3014/3015 (CW), and American Institutions (AI). The degree will also require 6 Intellectual Exploration (IE) courses, two Fine Arts (FF), two Humanities (HF), and two Social/Behavioral Science (BF). At least two must be upper division, at least one must satisfy the Diversity requirement, and at least one must satisfy the International requirements.

There is one required course in this category; namely, the new course
CS 4962, Ethics and Data Science
We are in the process of petitioning for this to satisfy the IE

and BF requirement. As this was first taught in Fall 2017, this is still a temporary number, so it is expected that the course number will change, but remain upper division.

3. Analytical Foundations (17-18 credit hours)
This component consists of the mathematical foundations required for more advanced data analysis, as well as some foundational parts of data analysis. CS 2100 is a computer science course that provides the combinatorial foundations necessary to understand many discrete processes and the scalability of algorithms. Math 2270 and Math 3070 (or CS 3130) are in linear algebra and probability which are the core mathematical foundations for most modern advanced data analysis. Math 3080 and CS 3190 are introductory views of the two main approaches to data analysis, statistical and optimization/cross-validation.

The set of probability and statistics courses – Math 3070, Math 3080, and CS 3130/ECE3530 (Engineering Probability and Statistics) – allows for some flexibility. The Math 3070 / Math 3080 sequence provides a broader and more mathematical view of statistical data analysis. While the CS 3130/ECE3530 course is designed as a more computational view of statistics, and by itself may be slightly more appropriate for a major in data science, but does not cover as much on the statistics side as the math sequence does. Initially the degree requires Math 3080 and either of Math 3070 or CS 3130/ECE3530, however, as demand exists (currently CS 3130 / ECE 3530 attracts more than 100 students a semester and Math 3070 and Math 3080 have multiple sections) there may be created a new course sequence

DS 3170 Applied Probability
DS 3180 Applied Statistics

which are more tailored to the needs to data science focused students. This would be a joint effort between the Computer Science and Mathematics departments.

4. Computing Foundations (10 credit hours)
The courses in this component provides the computer science and programming skills necessary to work as part of a software engineering team in designing and building reliable data processing and analysis software. CS 3300, Data Wrangling will be a new introductory course on parsing and manipulating data from various sources, preparing them so they can be analyzed.

CS 3500 is the first main software engineering course for Computer Science majors; it teaches how to manage large software projects, work in teams, and requires the completion of larger software programming assignments. CS 4150 teaches the design of scalable computational tools, so as large data sets are encountered the programs are ensured to run efficiently.

The CS 3300 course does not yet exist, but has been in discussion for a couple years. The SoC is committed to teaching it annually in the future, and MDCRC 6521 / BMI 6018 could potentially be used as a temporary replacement. Since it does not exist, it will be possible to build in various components not required elsewhere such as experimental design and human computer interaction.
With sufficient demand, SoC may redesign the CS 3500 course into a data science focused version
CS 3500, Data Science Projects
where the same software engineering and large project teamwork skills are taught, but the larger software projects are explicitly focused on data science topics.

5. Core Data Science (12 credit hours)
This component includes the 4 core courses which make this degree distinct, and are included in almost all undergraduate data science degrees offered elsewhere (see discussion on related degrees above). Across these courses (CS 5140, CS 5350, CS 5530, CS 5630) students learn advanced techniques on how to analyze, process, management, and communicate large, diverse, and heterogeneous data sets. These course each have ethical components built in, and so students are walked through and made aware of the ethical challenges and implications of these sorts of tasks.
While these course are (in the case of CS 5140, CS 5350, CS 5630) co-listed with graduate level course, undergraduates have for the last few years comprise about one quarter to one third of the students. Moreover, with the more developed preparatory courses (CS 3300, CS 3190), the students should be prepared for these courses.

6. Elective Advanced Data Analysis (9 credit hours)
Students will have flexibility to take advanced data science

courses, in mathematics of analysis (e.g., Math 5080, Math 5090), in computational analysis (e.g., CS 4300, CS 5190, CS 5340, CS 5640), or in other computational aspects related to data science (e.g., CS 5530, CS 6630, CS 5150).

7. Elective Data Domain (9-12 credit hours)
Students will engage in some domain which generates interesting data which is in need of management and analysis. They can choose 3 courses approved by program director. We have a partial list of pre-approved course sequences, including ones from Economics, Geography, and Atmospheric Sciences.

8. Capstone Project (6 credit hours)
The students will complete a capstone project in their senior year. This will follow the two course sequence found effective in Computer Science.

Degree Map

Degree maps pertain to undergraduate programs ONLY. Provide a degree map for proposed program. Degree Maps were approved by the State Board of Regents on July 17, 2014 as a degree completion measure. Degree maps or graduation plans are a suggested semester-by-semester class schedule that includes prefix, number, title, and semester hours. For more details see http://higheredutah.org/pdf/agendas/201407/TAB%20A%202014-7-18.pdf (Item #3).

Please cut-and-paste the degree map or manually enter the degree map in the table below.

The schedule below satisfies all degree and general education elective requirements, and constitutes 122 credits. The various required electives are marked accordingly:

[[Elective]] indicates an example general elective

[Elective (DD)] indicates an example data domain elective

[Elective (ADA)] indicates an example advanced data analysis elective

This schedule shows a path where a student tests out of CS 1030 (which is common for Computer Science majors). Alternatively if CS 1030 is taken in the first semester, then this can be accommodated, for example, by shifting CS 1410, CS 2420, CS 3500, and CS 5350 back one semester (these courses are offered every semester) and CS 4964 and CS 5630 back 2 semesters (these courses are currently only offered in Fall or Spring, respectively).

First Year Fall	Cr. Hr.	First Year Spring	Cr. Hr.
CS 1410 : Object-Orient. Prog	4	CS 2100 : Discrete Structures	3
Math 1310 : Engineering Calculus I	4	CS 2420 : Algorithms/Data Struct.	4
[[Writing Elective]]	3	Math 1320 : Engineering Calculus II	4
[Geo 1110 : Intro to Earth Systems (DD)]	3	[[Writing Elective]]	3
		[[Elective]]	3
Total	14	**Total**	17
Second Year Fall	**Cr. Hr.**	**Second Year Spring**	**Cr. Hr.**
CS 3500 : Software Practice	4	Math 2270 : Linear Algebra	4
CS 4964 : Foundations of Data Analysis	3	Math 3080 : Applied Statistics II	4
Math 3070 : Applied Statistics I	4	CS 3300 : Data Wrangling	3
CS 4962 : Ethics and Data Science	3	[GEO 3070 : Petrology for Engineers (DD)]	3
[Geo 3060 : Structural Geo and Tectonics (DD)]	3		
Total	17	**Total**	14
Third Year Fall	**Cr. Hr.**	**Third Year Spring**	**Cr. Hr.**
CS 4150 : Algorithms	3	CS 5140 : Data Mining	3
CS 5350 : Machine Learning	3	CS 5630 : Database Systems	3
[[Fine Arts Elective]]	3	[CS 4300 : Artificial Intelligence (ADA)]	3
[[Elective]]	3	[[Fine Arts Elective]]	3
[[Elective]]	3	[[Humanities Elective]]	3
Total	15	**Total**	15
Fourth Year Fall	**Cr. Hr.**	**Fourth Year Spring**	**Cr. Hr.**
CS 4000 : Senior Capstone Design	3	CS 4500 : Senior Capstone Project	3
CS 5630 : Visualization for Data Science	3	[Math 5090 : Statistical Inference II (ADA)]	3
[Math 5080 : Statistical Inference I (ADA)]	3	[[Social/Behavior Elective]]	3
[[Humanities Elective]]	3	[[American Institutions]]	3
[[Elective]]	3	[[Elective]]	3
Total	15	**Total**	15

Appendix C: Current and New Faculty / Staff Information

Part I. Department Faculty / Staff

Identify # of department faculty / staff (headcount) for the year preceding implementation of proposed program.

	# Tenured	# Tenure -Track	# Non -Tenure Track
Faculty: Full Time with Doctorate	27	15	7
Faculty: Part Time with Doctorate			
Faculty: Full Time with Masters			
Faculty: Part Time with Masters			
Faculty: Full Time with Baccalaureate			
Faculty: Part Time with Baccalaureate			
Teaching / Graduate Assistants			
Staff: Full Time			14
Staff: Part Time			3

Part II. Proposed Program Faculty Profiles

List current faculty within the institution -- with academic qualifications -- to be used in support of the proposed program(s).

	First Name	Last Name	Tenure (T) / Tenure Track (TT) / Other	Degree	Institution where Credential was Earned	Est. % of time faculty member will dedicate to proposed program.	If "Other," describe
Full Time Faculty							
	Aditya	Bhaskara	TT	PhD	Princeton University	50	
	Elaine	Cohen	T	PhD	Syracuse University	25	
	Tom	Henderson	T	PhD	University of Texas	25	
	Tucker	Hermans	TT	PhD	Georgia Tech	25	
	John	Hollerbach	T	PhD	MIT	25	
	Alexander	Lex	TT	PhD	Graz University of Technology	50	
	Feifei	Li	T	PhD	Boston University	50	
	Valerio	Pascucci	T	PhD	Purdue University	25	
	Jeff	Phillips	T	PhD	Duke University	100	
	Srikumar	Ramalingam	TT	PhD	INRIA Grenoble Rhone-Alpes	25	
	Ellen	Riloff	T	PhD	University of Massachusetts	25	
	Vivek	Srrikumar	TT	PhD	University of Illinois	50	
	Suresh	Venkatasubramanian	T	PhD	Stanford University	50	
	Bei	Wang Phillips	TT	PhD	Duke University	25	
	Ross	Whitaker	T	PhD	University of North Carolina	25	
	Jason	Wiese	TT	PhD	Carnegie Mellon University	25	
	Shandian	Zhe	TT	PhD	Purdue University	50	
Part Time Faculty							

	First Name	Last Name	Tenure (T) / Tenure Track (TT) / Other	Degree	Institution where Credential was Earned	Est. % of time faculty member will dedicate to proposed program.	If "Other," describe

Part III: New Faculty / Staff Projections for Proposed Program

Indicate the number of faculty / staff to be hired in the first three years of the program, if applicable. Include additional cost for these faculty / staff members in Appendix D.

	# Tenured	# Tenure -Track	# Non -Tenure Track	Academic or Industry Credentials Needed	Est. % of time to be dedicated to proposed program.
Faculty: Full Time with Doctorate					
Faculty: Part Time with Doctorate					
Faculty: Full Time with Masters					
Faculty: Part Time with Masters					
Faculty: Full Time with Baccalaureate					
Faculty: Part Time with Baccalaureate					
Teaching / Graduate Assistants					
Staff: Full Time					
Staff: Part Time					

Appendix D: Projected Program Participation and Finance

Part I.

Project the number of students who will be attracted to the proposed program as well as increased expenses, if any. Include new faculty & staff as described in Appendix C.

Three Year Projection: Program Participation and Department Budget						
	Year Preceding Implementation	New Program				
		Year 1	Year 2	Year 3	Year 4	Year 5
Student Data						
# of Majors in Department	561	590	620	640	660	680
# of Majors in Proposed Program(s)		10	25	40	60	70
# of Graduates from Department	128	130	140	145	150	155
# Graduates in New Program(s)		0	5	10	25	30

Department Financial Data				
	Department Budget			
		Year 1	Year 2	Year 3
Project additional expenses associated with offering new program(s). Account for New Faculty as stated in Appendix C, "Faculty Projections."	Year Preceding Implementation (Base Budget)	Addition to Base Budget for New Program(s)	Addition to Base Budget for New Program(s)	Addition to Base Budget for New Program(s)
EXPENSES – nature of additional costs required for proposed program(s)				
List salary benefits for additional faculty/staff each year the positions will be filled. For example, if hiring faculty in year 2, include expense in years 2 and 3. List one-time operating expenses only in the year expended.				
Personnel (Faculty & Staff Salary & Benefits)				
Operating Expenses (equipment, travel, resources)				
Other:				
TOTAL PROGRAM EXPENSES		$0	$0	$0
TOTAL EXPENSES	$0	$0	$0	$0
FUNDING – source of funding to cover additional costs generated by proposed program(s)				
Describe internal reallocation using Narrative 1 on the following page. Describe new sources of funding using Narrative 2.				
Internal Reallocation				
Appropriation				
Special Legislative Appropriation				
Grants and Contracts				
Special Fees				
Tuition				
Differential Tuition (requires Regents approval)				
PROPOSED PROGRAM FUNDING		$0	$0	$0
TOTAL DEPARTMENT FUNDING	$0	$0	$0	$0
Difference				
Funding - Expense	$0	$0	$0	$0

Part II: Expense explanation

Expense Narrative

Describe expenses associated with the proposed program.

We do not expect any financial expenses due to this program.

Part III: Describe funding sources

Revenue Narrative 1

Describe what internal reallocations, if applicable, are available and any impact to existing programs or services.

We do not plan any internal reallocations, since no new expenses are planned.

Revenue Narrative 2

Describe new funding sources and plans to acquire the funds.

We do not plan to seek any new funding sources since no new funding is expected to be needed.