UNIVERSITY OF GEORGIA

University of Georgia
Athens, Georgia 30602
univcouncil@uga.edu
www.uga.edu

University Council

August 20, 2021

UNIVERSITY CURRICULUM COMMITTEE – 2021-2022
Susan Sanchez, Chair
Agricultural and Environmental Sciences – Nicholas Fuhrman
Arts and Sciences – Jonathan Haddad (Arts)
Rodney Mauricio (Sciences)
Business – Jim Carson
Ecology – Amanda Rugenski
Education – David Jackson
Engineering – Kun Yao
Environment and Design – Ashley Steffens
Family and Consumer Sciences – Sheri Worthy
Forestry and Natural Resources – Joseph Dahlen
Journalism and Mass Communication – Dodie Cantrell-Bickley
Law – Randy Beck
Pharmacy – Michelle McElhannon
Public and International Affairs – Leah Carmichael
Public Health – Allan Tate
Social Work – Harold Briggs
Veterinary Medicine – Shannon Hostetter
Graduate School – Wendy Ruona
Ex-Officio –Provost S. Jack Hu
Undergraduate Student Representative – Matthew Jue
Graduate Student Representative – TBD

Dear Colleagues:

The attached proposal from the Franklin College of Arts and Sciences for a new major in Data Science (M.S.) will be an agenda item for the August 27, 2021, Full University Curriculum Committee meeting.

Sincerely,

Susan Sanchez, Chair
University Curriculum Committee

cc: Provost S. Jack Hu
Dr. Rahul Shrivastav

USG ACADEMIC PROGRAM PROPOSAL
(Effective 2/22/18)

Institution: The University of Georgia

Date Completed at the Institution: November 24, 2020

Name of Proposed Program/Inscription: Data Science (M.S.)

Degree: Master of Science (M.S.)

Major: Data Science

CIP Code: 30300101

School/Division/College: Franklin College of Arts and Sciences

Department: Computer Science, Statistics

Anticipated Implementation Date: Spring 2022

Requesting Differential Tuition Rate: No

Delivery Mode (Check all that apply):

On-campus, face-to-face only	**X**
Off-campus location, face-to-face only (specify the location):	
Online Only	
Combination of on-campus and online (specify whether 50% or more is offered online for SACSCOC)	
Combination of off-campus and online (specify whether 50% or more is offered online for SACSCOC)	
Hybrid, combination delivery, but less than 50% of the total program is online based on SACSCOC	
Contractual Location (specify the location and timeframe/start and end dates):	

[1] **All documents and forms requesting a differential tuition rate must be submitted to the Office of Fiscal Affairs prior to Academic Affairs Review of the Degree Proposal.**

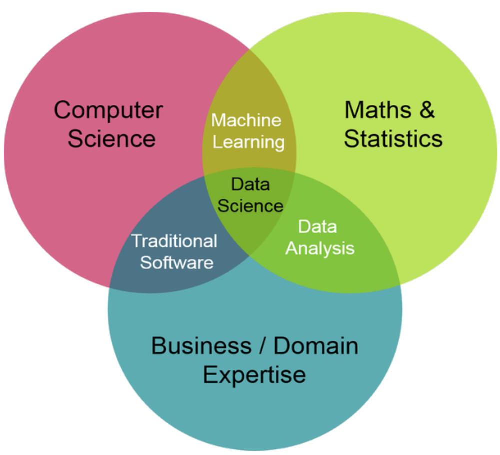
1) **Forecast:** If this program was not listed on one of the past two-year academic forecasts, provide an explanation concerning why it was not forecasted but is submitted at this time.

This program was not included in the University of Georgia's Academic Forecast because it had not been submitted through the faculty governance process.

2) **Academic Framework:** Within the context of strategic planning of all resources and divisions within short-term and long-term perspectives, provide a narrative that explains campus leadership review and attention to newly institutionally approved programs within the last four years, low-producing programs, and post-approval enrollment analyses prior to approving the proposed program for submission to the system office.

The Office of Instruction reviews newly institutionally-approved programs, low-producing programs, and post-approval enrollment to monitor and assess future viability of all programs.

3) **Rationale:** We now live in the "Big Data" era: rapid advances in computer processing power, networking speed, and storage capacity; large-scale and real-time data collection in science, business, industry, and government. This perfect storm of data availability and computational power has given rise to "Data Science" emerging as a prominent field of study for its ability to provide valuable insights for making informed business, health, or scientific decisions. The field of Data Science was in fact theorized over 50 years ago by John Tukey in his 1962 book, *The Future of Data Analytics*, in which he presented the broad topics of data analytics, interpretation, and visualization (Fig. 1) as their own field, rather than extensions or branches of mathematics and statistics. He argued that there is considerable value in training students in the practice of extracting information from data.



Data Science encompasses a wide range of concepts, methodologies, and algorithms involved in the collection, management, visualization, analysis, and transformation of data into one of many possible end products: precision information, event inference and prediction, knowledge creation, or decision making. The word *Data* refers to data collection, storage, management, and retrieval, while the word *Science* deals with modeling, analysis, inference, interpretation, and decision making. Inherently, Data Science is a field of study that rests at the intersection of Computer Science and Statistics. However, it also emphasizes interdisciplinarity and bridge-building, given the applicability of Data Science principles and practices to countless fields outside of Computer Science and Statistics.

Voluminous data is being regularly collected and analyzed in science, business, and industry, as well as by government and society at large. This "data deluge" is fundamentally changing the way corporations do business and is also leading to new discoveries in every scientific field. The collection, management, analysis, and interpretation of such data, with complex structures in the form of text, video, and streaming data, are paving the way to exciting new research avenues. The U.S. government and industry are investing substantial resources in research and development under the umbrella of "Data Science" and its most visible subfield, "Artificial Intelligence."

Currently, the Computer Science department offers courses related to Data Science and an undergraduate Certificate of Applied Data Science program. However, this certificate only requires minimal coursework in mathematics and statistics. These fundamental topics require substantially more instruction time for a rigorous Data Science program, from designing computational models in different fields to studying advanced topics in data analytics. The Terry College of Business offers a one-year (two-semester) Master of Science in Business Analytics. However, the Business Analytics program is aimed only at graduate students who are primarily interested in business applications. The Computer Science and Statistics departments recently launched a joint undergraduate program in Data Science, to great success, but this has once again left out graduate students in the computer and mathematical sciences who wish to study the research domains of Data Science. Thus, at the University of Georgia, there is no comprehensive graduate master's program that adequately trains students in all the areas involved in Data Science.

The departments explored the idea of offering Data Science as a track in the Statistics (M.S.) program or choosing a single home department for the Data Science (M.S.) program: either Computer Science or Statistics. Neither model was pursued because, as noted earlier, Data Science is a field that sits squarely between Computer Science and Statistics, and the success of the program will require the support and full investment from both groups of faculty.

The proposed Data Science (M.S.) degree will bridge key elements of Computer Science and Statistics into an interdisciplinary degree to develop future generations of data scientists. Faculty members in Computer Science and Statistics have spent more than a year envisioning, planning, and designing the new program. It has the strong support of the dean of Franklin College, heads of the two departments, and the Graduate Program Faculty members in both departments. In addition, the proposal also includes a strong letter of support from the dean of Terry College of Business.

4) **Mission Fit and Disciplinary Trends:** Description of the program's fit with the institutional mission and nationally accepted trends in the discipline (explain in narrative form). If the program is outside the scope of the institutional mission and sector, provide the compelling rationale for submission.

The Master of Science in Data Science will fit the mission of the University of Georgia as it provides the necessary expertise of graduates in the high-demand area of data science (including data management, data analytics, and machine learning). One of the missions of UGA is its commitment to excellence in public service, economic development, and technical assistance activities designed to address the strategic needs of the state of Georgia. This program will support this mission of UGA by providing a well- trained workforce in the aforementioned data science area.

The development of this M.S. program fits in well with our larger goal of establishing the University of Georgia as a leader in the field of Data Science. The proposed Data Science (M.S.) degree will build a strong relationship with the Georgia Informatics Institutes for Research and Education (GII) at UGA, which is serving as a hub for informatics research and instruction. Master's students in the Data Science program will benefit greatly from the synergistic activities organized by GII including working with GII-faculty on research projects through directed studies.

This M.S. program also fits in with the trends in the nation. In order to meet the immediate demand for data scientists, many universities across the U.S. have launched Data Science programs, including University of Virginia, University of Wisconsin, Carnegie Mellon University, Columbia University, Harvard University, University of Missouri, University of Texas at Austin, Syracuse University, University of Denver, Southern Methodist University, and Duke University, some of which are fully online.

5) **Description and Objectives:** Program description and objectives (explain in narrative form).

The demand for data scientists is increasing rapidly as large-scale data are being heavily utilized by industries and scientific research. Progress in Statistics, Data Mining, and Machine Learning combined with greater computational capabilities has resulted in dramatic improvements in the ability to analyze both structured (e.g., tabular) and unstructured (e.g., text, audio, image, and video) data. A data scientist will need to be able to manage the full lifecycle of data, which requires knowledge and skills from both Computer Science and Statistics. The proposed graduate program in Data Science will provide the students with a strong foundation in Data Science, covering algorithms, distributed systems, database management, and machine learning (in Computer Science), and regression, time-series analysis, design of experiments, statistical learning, and Bayesian statistics (in Statistics). Students graduating with an M.S. in Data Science will know how to develop software, design and maintain databases, process data in distributed environments, analyze the data using techniques from statistics, data mining and machine learning, provide visualizations of the data or the results of analysis, and assist decision makers. The program will include practical application of acquired knowledge and skills in the form of a Master's Project course. Upon graduation, students will be in high demand in the workforce (e.g., Google, Amazon, Facebook, Coca-Cola, UPS, Delta Airlines, Home Depot, IBM, Intel, Samsung, Boeing, Goldman Sachs, AIG, Liberty Mutual, Johnson & Johnson, NASA, NIST, DoD) or continue their education at the doctoral level. Several academic avenues should also be open to students, including Ph.D. in Data Science, Computer Science, Statistics, Management Information Systems, or Industrial Engineering.

6) **Need:** Description of the justification of need for the program. (Explain in narrative form why the program is required to expand academic offerings at the institution, the data to provide graduates for the workforce, and/or the data in response to specific agency and/or corporation requests in the local or regional area, and/or needs of regional employers.) (A list of resources, not exhaustive, is available on the public web link along with the proposal form at: http://www.usg.edu/academic_programs/new_programs)

As indicated in John Tukey's book *Exploratory Data Analysis*, the concept of a field of science dedicated to the curation, interpretation, and visualization of data is not novel to the current generation. Rather, a confluence of very specific events unique thus far to our society have unlocked the potential for Data Science as a field. First and foremost, the effects of the internet have been felt across every aspect of our lives, and research is no exception. The ability to share data, collaboratively design methods, and engage in real-time across vast distances has fundamentally changed, and will continue to change, how we conduct and reproduce original research. Second, the abundance of cheap, dense digital storage mechanisms has fundamentally altered how datasets are created and stored. This has resulted in an explosion of digital datasets and led to the rise of the term "Big Data" that is now so ubiquitous. Huge amounts of data are being collected in all areas, made possible by rapid advances in computing, measurement, data storage, and data transfer technologies. Examples in business and industry range from transactional data captured by companies via the internet and social media to sensor data captured by smartphones, automobiles, industrial systems, and environmental networks. Examples in government include monitoring weather and environmental systems, social and health services, regulatory compliance, and national security. Examples in science, including the Sloan Digital Sky Survey, the Great Lakes Environmental Database, the GenBank DNA sequence database, etc., are driving scientific discovery and creation of new knowledge. Precision agriculture is but one area that straddles all three of these broad categories. Put simply, the enormous interest in Data Science now is a direct consequence of this confluence.

In 2012, the *Harvard Business Review* reported that the "sexiest job of the 21st century" was that of a data scientist. In January 2019, *Forbes* reported that Glassdoor named data scientist the best job in America for the 4th consecutive year. As of March 2020, Glassdoor had 19,357 data scientist positions posted. The Brookings Institution's *Tech Stream* is reporting how data science can help ease the COVID-19 pandemic, while one of UNICEF's data bosses tells *Forbes* about the need for rapid data collection to identify the needs of the acutely vulnerable, and *The Economist* discusses managing pandemics with data. The need for data scientists changes with the times, but as long as data is collected in ways that continue to push the boundaries of "enormous," as with the Internet of Things, the need for data scientists will persist.

Big Data and Data Science are continuing to be transformative in all areas of inquiry as the amount of data collected and stored increases from about 40 zettabytes in 2019 to 175 zettabytes in 2025. As such, students graduating with a degree in Data Science will have a huge range of job opportunities. According to the Bureau of Labor Statistics Occupational Outlook Handbook, for data scientists and mathematical science occupations, the 2019 number employed is 33,200 and the May 2019 median annual wage is $94,280. Furthermore, from 2019 to 2029, the projected number of new jobs is 10,300 and the employment of data scientists is projected to grow 31% (much faster than average) during the 10-year timeframe. In fact, according to the Georgia Bureau of Labor Statistics Occupational Employment Statistics, for data scientists and mathematical science occupations, Georgia is among the top 10 states in the U.S. and metro-Atlanta is among the top 10 metropolitan areas in the U.S. with the highest employment levels.

	Employment	Location Quotient	Employment per 1000	Annual Mean Wage
U.S.	30,810	--	0.21	$100,560
Georgia	1,270	1.35	0.28	$81,520
Atlanta	920	1.60	0.34	$90,040

Since 2007, when North Carolina (NC) State University first established a Master of Science in Analytics (M.S.A.), the job market for people with training in any sort of data science/analysis/analytics has exploded; according to *Business Insider*, it is among the most recession-proof of master's degrees. The 2019 Burtch Work Study reports a median salary for Data Science level 1 professionals of $95,000, with Data Science managers earning well over $145,000 annually.

Many universities have established or are beginning to establish both graduate and undergraduate programs in Data Science – anecdotally, there now exists over 300 such programs across the U.S. The site mastersindatascience.org lists what they have determined to be the best 23 schools offering a master's degree in Data Science, while the Data Science Degree Programs Guide gives a list of its 30 best master's programs in Data Science.

The University of Georgia's major in Data Science (B.S.) just accepted the first cohort of students in fall 2019. The table below gives the enrollment figures in all Data Science programs at UGA; also shown are the enrollments for the first such program in the U.S. at NC State for its inaugural year and for the two most recent years. These figures serve to indicate what is possible locally, as well as indicate the potential for growth over the long run.

School	Program	Unit	Year	Applications	Admitted	Enrolled	Graduated
NC State	MS in Analytics	Institute for Advanced Analytics	2020-21	1028	154	124	--
			2019-20	1131	151	119	116
			2008-09	48	35	24	22
UGA	MS in Business Analytics	Terry College of Business	2020-21	186	93	56	--
			2019-20	130	43	36	36
			2018-19 (inaugural)	91	33	25	24
	BS in Data Science	Dept of Statistics	2020-21	NA	NA	71	--
			2019-20 (inaugural)	NA	NA	28	--
	Certificate in Applied Data Science	Dept of Computer Science	2020-21	NA	NA	111	--
			2019-20	NA	NA	105	16
			2018-19	NA	NA	88	5
			2017-18	NA	NA	50	0
			2016-17	NA	NA	11	0

This new program will further enhance existing collaborations between the Computer Science and Statistics departments spearheading the program. There is considerable overlap in research interests in faculty and students in the two departments – in computational biology, machine learning, biomedical informatics, bioinformatics, and data-intensive computational techniques, for example. This program will also serve as a focal point for other units on campus to offer related courses, and as a clearinghouse for research opportunities, internships, and jobs.

Waiting to develop this program brings significant risks. There are societal risks of not being able to meet the significant projected demand in the field for data analysts with deep Data Science skills in the coming decade and falling behind other institutions. Even with all of the programs in existence across the US, there are not enough skilled Data Scientists to meet the ever-increasing demand. Companies in Atlanta alone have a large and growing appetite for employees with Data Science skills. This program will make a big impact in the available Data Science talent for the entire State of Georgia.

7) **Demand:** Please describe the demand for the proposed program. Include in this description the supporting data from 1) existing and potential students and 2) requests from regional industries. How does the program of study meet student needs and employer requirements in terms of career readiness and employability, requirements to enter the profession, post-graduate study, and disciplinary rigor at the level required for professional success and advanced educational pursuits? *(In other words, how does the program of study prepare students for the next step?)*

Data Science is already a fast-growing area and there is more growth expected over the next few decades. Enormous data generation in research, business, government, and society is fueling the necessity for highly-trained data scientists who can manage, manipulate, and model voluminous data. It is well known that the demand for data scientists in the workforce over the next 5 to 10 years will far outpace supply.

For four years in a row, data scientist has been named the number one job in the U.S. by Glassdoor. What's more, the U.S. Bureau of Labor Statistics reports that the demand for data science skills will drive a 27.9 percent rise in employment in the field through 2026. Not only is there a huge demand, but there is also a noticeable shortage of qualified data scientists (https://data-flair.training/blogs/data-science-job-trends/).

The huge demand for professionals skilled in data, analytics, machine learning, and AI will be addressed by the education sector. Many industries, such as the cybersecurity industry, the aviation industry, healthcare industries, agriculture, and many others, are investing heavily in data science approaches to address their needs.

All of those sectors are trying to get their hands on the limited number of data scientists available. Since 39% of the most rigorous data science positions require a degree higher than a bachelor's degree, the whole market, and the education sector in particular, will be playing catch-up from now on. Until there are enough highly skilled professionals for all these roles, companies will have to resort to creative solutions to fill those gaps (https://www.bmmagazine.co.uk/business/data-science-will-grow-in-scope-by-2020/).

The proposed M.S. program offered jointly between the Computer Science and Statistics Departments will provide the graduates with the analytical skills that are needed for Data Scientists that include the statistics and computer science rigorous advanced topics (https://www.morningfuture.com/en/article/2018/02/21/data-analyst-data-scientist-big-data-work/235/).

The Department of Computer Science has more than 1,155 undergraduate majors and more than 120 graduate students who pursue a master's degree in Computer Science each year. The Department of Statistics has about 180 undergraduate Statistics majors and about 50 students pursuing a master's degree in Statistics each year. At least 30 of the 50 graduate students who earn an M.S. degree in Statistics are dual-degree students who are simultaneously earning an M.S. or a Ph.D. degree in another discipline within UGA. The number of undergraduate Data Science majors has grown tremendously from 27 in fall 2019 to 56 in fall 2020. Therefore, the pool of potential students for the proposed M.S. in Data Science degree, even within UGA, is rather large since it would consist of students currently pursuing a B.S. in Computer Science, Statistics, or Data Science, and some of those who wish to pursue an M.S. degree in Computer Science or Statistics. Furthermore, many of the Computer Science and Statistics courses listed in the curriculum of the proposed M.S. program are popular on campus and have been experiencing a steady increase in enrollments.

At the national level, a large number of universities have started graduate programs in Data Science, and their programs are experiencing an undiminished and sustained upward trend. To gauge the interest among UGA students in pursuing the proposed M.S. degree in Data Science, on June 15, 2020, the departments conducted a formal, anonymous survey of UGA undergraduate Computer Science and Statistics majors who have completed at least 20 credit hours of courses in either discipline. The students were asked the following: "If an M.S. degree in Data Science was available in the Computer Science and Statistics Department at UGA next year, please indicate your level of interest in pursuing such degree: 5: No Interest, 4: Not sure, 3: Would consider, 4. Probably Yes, 1: Definitely Yes. Please circle only one choice."

Of the 70 students who responded to the survey, 16 picked Definitely Yes, 18 selected Probably Yes, 26 Would Consider, 6 Not Sure, and 4 No Interest. The same survey was sent to the students enrolled in the B.S. in the Data Science degree program that started in fall 2019. Of the 20 students who responded to the survey, 4 said Definitely Yes, 6 selected Probably Yes, 7 Would Consider, 3 Not Sure, and 0 No Interest. The low response rate was probably due to the pandemic. Nonetheless, a strong demand is expected for such a degree.

For the first year, the projected enrollment in the proposed M.S. in Data Science is approximately 15 students. Each year thereafter, the program is expected to grow by at least five new students. The B.S. in Data Science degree was introduced in fall 2019 with an enrollment of 27 students, and has grown to about 56 students in fall 2020. The total number of students in the master's program is expected to grow to about 52 students by 2024. In the long term, the program is expected to have a healthy growth, given the demand for an M.S. in Data Science degree in the nation.

An important feature of this major will be its likely appeal to a diverse student body. As is well known, female students are seriously underrepresented in mathematics and computer science. However, enrollment of female undergraduate students in Statistics more closely mirror the overall campus gender distribution. Statistics is also beginning to see underrepresented students pursuing an M.S. in Statistics through the Graduate Feeder Program at UGA. The departments will send informative flyers about the M.S. in Data Science program to public and private schools, including Historically Black Colleges, in Georgia and neighboring states. This master's program will be a very attractive as a UGA Double Dawgs programs such as a B.S. in Computer Science and M.S. in Data Science, B.S. in Statistics and M.S. in Data Science, B. S. in Engineering and M.S.in Data Science, B.S. in Data Science and M.S. in Data Science, and B.S. in Mathematics and M.S. in Data Science.

8) **Duplication:** Description of how the program does not present duplication of existing academic offerings in the geographic area, within the system as a whole, and within the proposing institution regardless of academic unit. If similar programs exist, indicate why these existing programs are not sufficient to address need and demand in the state/institution's service region and how the proposed program is demonstrably different or complementary to other USG degrees and majors.

Currently, there is no master's degree in Data Science offered at the University of Georgia. The following table gives a list of seven (7) related master's level programs that are (or will be) offered at different public schools in Georgia.

Institution Name	Title of the Program with Link	Notes
University of Georgia	Graduate Certificate in Agricultural Data Science	The College of Agricultural and Environmental Sciences offers this program
University of Georgia	Business Analytics (M.S.)	The Terry College of Business offers this one-year program.
Georgia Institute of Technology	Analytics (M.S.)	The College of Business, the College of Computing, and the College of Engineering this program. This is also offered online.
Georgia State University	Data Science and Analytics (M.S.)	The College of Business offers this program.
Mercer University	Business Analytics (M.S.)	The School of Business offers this program.
Southwest Georgia Technical College	Predictive Analytics (M.S.)	This online program is offered by the Moravian Graduate Online.
Augusta University	Data Science (M.S.)	This program will be launched in fall 2021. It is

		jointly administered by the Division of Biostatistics and Data Science in the Department of Population Health Sciences.

All the master's programs listed in the table above, except the yet to be launched Data Science (M.S.) at August University, have a strong focus toward business applications since the programs are either jointly or solely administered by respective business schools. The program at Augusta University focuses on applications in the health sciences. In contrast, the proposed Data Science (M.S.) is unique in that it will be jointly administered by two strong disciplines—Computer Science and Statistics—which are the foundational pillars of data science. This partnership between the two disciplines makes this program demonstrably unique in the state of Georgia in terms of foundational training, breadth of application, and addressing the global need and demand for data scientists with deep analytical skills. In fact, unlike the programs listed in the table, the program of study in the proposed Data Science (M.S.) places equal emphasis on both computer science and statistics in terms of the core courses while allowing the students to select from a large array of elective courses in both disciplines to customize their focus. The proposed program is designed to produce the next generation of data scientists who can not only construct novel statistical models to analyze big data arising in or outside of a business/health context, but also create computational methodologies and tools to effectively leverage the available big data, process them, and draw statistically valid conclusions. Thus, the proposed Data Science (M.S.) does not present a duplication of the existing master's programs within Georgia. More importantly, given the high demand for data scientists in state of Georgia, even with the yet to be launched Data Science (M.S.) program at Augusta University, there will not be enough skilled Data Scientists to meet the ever-increasing demand. Therefore, the proposed program has the potential to make a big impact in the available Data Science talent for the entire State of Georgia.

Finally, another unique aspect that distinguishes the proposed Data Science (M.S.) from the other programs in Georgia is that it will provide a natural pathway for UGA's undergraduate Data Science majors and provide an opportunity to create a UGA Double Dawgs pathway to ambitious and motivated students who want to earn both a bachelor's degree and a master's degree in five years or less, thereby saving time and money while positioning themselves for a successful career in data science after graduation.

9) **Collaboration:** Is the program in collaboration with another USG Institution, TCSG institution, private college or university, or other entity?

Yes ___ or No __**X**__ (place an X beside one)

If yes, list the institution below and include a letter of support from the collaborating institution's leadership (i.e., President or Provost and Vice President for Academic Affairs) for the proposed academic program in Appendix I.

10) Admission Criteria: List the admission criteria for the academic program, including standardized test and grade point average requirements for admission into the program. Also, at what point (e.g., credit hours completed) are students admitted to the program.

In addition to the general University of Georgia requirements and policies set forth in the Graduate Bulletin, the following requirements and policies apply to all applicants:

1. A bachelor's degree is required, preferably with a major in a STEM+Computing field. Students with insufficient background in Computer Science or Statistics must take remedial Computer Science/Statistics courses, in addition to their graduate program, to address any deficiencies. A sufficient background for the MS in Data Science must include at least the following courses (or their equivalent):

MATH 2250, Calculus I for Science and Engineering (Differential Calculus), 4 hours, prereq: MATH 1113
MATH 2260, Calculus II for Science and Engineering (Integral Calculus), 4 hours, prereq: MATH 2250 or MATH 2300H
MATH 2270, Calculus III for Science and Mathematics (Multivariable Calculus), 4 hours, prereq: MATH 2260 or MATH 2310H or MATH 2410 or MATH 2410H
MATH 3300, Applied Linear Algebra, 3 hours, prereq: MATH 2260 or MATH 2410 or MATH 2410H
CSCI 1301-1301L, Introduction to Computing and Programming, 4 hours, prereq: MATH 1113
CSCI 1302, Software Development, 4 hours, prereq: CSCI 1301-1301L
CSCI 2150-2150L, Introduction to Computational Science, 4 hours, prereq: MATH 1113
CSCI 2725 Data Structures for Data Science, 4 hours, prereq: CSCI 1302
STAT 2010 Statistical Methods for Data Scientists, 3 hours, prereq: none

2. Admission to this program is selective; students with a record of academic excellence have a better chance of acceptance. Students with exceptionally strong undergraduate records may apply for admission to the graduate program prior to fulfilling all of the above requirements.

3. Graduate Record Examination (GRE) test scores are required for admission consideration. International applicants also need TOEFL, IELTS, or DuoLingo official test scores.

4. Three letters of recommendation are required, preferably written by university professors familiar with the student's academic work and potential. If the student has work experience, one letter may be from his/her supervisor.

5. A one or two-page personal statement outlining the student's background, achievements, and future goals is required.

6. A student may include a recent copy of his/her resume as part of the application packet; however, this is not required.

11) Curriculum

a. Specify whether the proposed program requires full-time study only, part-time study only, or can be completed either full time or part time.

This program of study can be completed either full-time or part-time.

b. If the proposed program will be offered online, describe measures taken by the academic unit to sufficiently deliver the program via distance education technologies and provide instructional and learning supports for both faculty and students in a virtual environment. Will the program be offered in an asynchronous or synchronous format?

This program will not be offered online.

c. List the entire course of study required to complete the academic program. Include the course prefixes, course numbers, course titles, and credit hour requirement for each course. Indicate the word "new" beside new courses. Include a program of study.

For the M.S. in Data Science degree, the table below lists the "**Required Courses**" under C.1, C.2, and C.3. a or b (at least 18 credit hours) and the "**Electives**" (at least 14 credit hours). In C.3, students are allowed to choose either a non-thesis (C.3.a) option or a thesis (C.3.b) option. Students must additionally complete CSCI 3030 or STAT 8920 if they have not already taken a suitable ethics course. The coursework consists of at least 32 semester hours.

	Course Prefixes & Numbers	**Course Titles**	**Credit Hours**	
	C.1. Core Courses			8
1	CSCI (STAT) 6375	Foundations of Data Science **(NEW)**	4	
2	CSCI 6360	Data Science II	4	
	C.2. Advanced Courses			6
	STAT 6420	Applied Linear Models	3	
3	or	or		
	STAT 6530	Statistical Inference for Data Scientists	3	
4	STAT 8330	Advanced Statistical Applications and Computing	3	
	C.3.a. Non-thesis option			4
	CSCI 7200	Masters Project	4	
	Or	or		

	STAT 7000	Master's Research		
5	**C.3.b. Thesis option** CSCI 7300 or STAT 7300	Master's Thesis or Master's Thesis	4	
	Electives (Two courses from each category)*			14*
6	Category A (see below)	CSCI Elective 1	4	
7		CSCI Elective 2	4	
8	Category B (see below)	STAT Elective 1	3	
9		STAT Elective 2	3	
		Total Credit Hours		**32**

*The **14 credit hours** of Electives mentioned in the table above will consist of **8 credit hours** from **Category A (Computer Science)** and **6 credit hours** from **Category B (Statistics)**:

Category A
CSCI 6150 (4 hours) - Numerical Simulations in Science and Engineering
CSCI 6170 (4 hours) - Introduction to Computational Investing
CSCI 6210 (4 hours) - Simulation and Modeling
CSCI 6370 (4 hours) - Database Management
CSCI 6380 (4 hours) - Data Mining
CSCI 6470 (4 hours) - Algorithms
CSCI 6780 (4 hours) - Distributed Computing Systems
CSCI 6795 (4 hours) - Cloud Computing
CSCI 6850 (4 hours) - Biomedical Image Analysis
CSCI 8360 (4 hours) - Data Science Practicum
CSCI 8370 (4 hours) - Advanced Database Systems
CSCI 8380 (4 hours) - Advanced Topics in Information Systems
CSCI 8535 (4 hours) - Multi Robot System
CSCI 8790 (4 hours) - Advanced Topics in Data Intensive Computing
CSCI 8820 (4 hours) - Computer Vision and Pattern Recognition
CSCI 8850 (4 hours) - Advanced Biomedical Image Analysis
CSCI 8920 (4 hours) - Decision Making Under Uncertainty
CSCI 8945 (4 hours) - Advanced Representation Learning
CSCI(ARTI) 8950 (4 hours) - Machine Learning
CSCI 8951 (4 hours) - Large-Scale Optimization for Machine Learning
CSCI 8955 (4 hours) - Advanced Data Analytics: Statistical Learning and Optimization.
CSCI 8960 (4 hours) - Privacy-Preserving Data Analysis

Category B:

STAT 6240 (3 hours) – Sampling and Survey Methods
STAT 6250 (3 hours) - Applied Multivariate Analysis and Statistical Learning
STAT 6280 (3 hours) - Applied Time Series Analysis
STAT 6350 (3 hours) - Applied Bayesian Statistics
STAT 6430 (3 hours) - Design and Analysis of Experiments
STAT 6510 (3 hours) - Mathematical Statistics I
STAT 6620 (3 hours) - Applied Categorical Data Analysis
STAT 6800 (3 hours) - Tools for Statistical Theory
STAT 8000 (3 hours) - Introductory Statistical Collaboration
STAT 8060 (3 hours) - Statistical Computing I
STAT 8070 (3 hours) - Statistical Computing II
STAT 8210 (3 hours) - Multivariate: Theory and Methods
STAT 8230 (3 hours) - Applied Nonlinear Regression
STAT 8260 (3 hours) - Theory of Linear Models
STAT 8270 (3 hours) - Spatial Statistics
STAT 8280 (3 hours) - Time Series Analysis
STAT 8290 (3 hours) - Advances in Experimental Designs
STAT 8620 (3 hours) - Categorical Data Analysis and Generalized Linear Models
STAT 8630 (3 hours) - Mixed-Effect Models and Longitudinal Data Analysis

d. State the total number of credit hours required to complete the program, but do not include orientation, freshman year experience, physical education, or health and wellness courses that are institutional requirements as defined in the Academic and Student Affairs Handbook, Section 2.3.1 and the Board Policy Manual, 3.8.1.

The program of study will require 32 credit hours to complete.

e. Within the appendix, append the course catalog descriptions for new courses and their prerequisite courses. Include the course prefixes, course numbers, course titles, and credit hour requirements.

Please see Appendix A.

f. If this is an undergraduate program, how does or would the department/institution use eCore, eMajor, or dual enrollment?

N/A

g. If this is a doctoral program, provide the names of four external reviewers of aspirational or comparative peer programs complete with name, title, institution, e-mail address, telephone number, and full mailing address. External reviewers must hold the rank of associate professor or higher in addition to other administrative titles.
N/A

12) Program of Study-Undergraduate Only

N/A

Program of Study- Graduate Only (Sample Program).

Courses (list acronym, number, and title)	**Semester**	**Hours**
CSCI(STAT) 6375, Foundations of Data Science (NEW)	First Year, Fall	4
STAT 6420, Applied Linear Models	First Year, Fall	3
STAT 6250, Applied Multivariate Analysis and Statistical Learning	First Year, Fall	3
CSCI 6360, Data Science II	First Year, Spring	4
CSCI 6370, Database Management	First Year, Spring	4
STAT 8000, Introductory Statistical Collaboration	First Year, Spring	3
CSCI 8360, Data Science Practicum	Second Year, Fall	4
CSCI 7200, Masters Project	Second Year, Fall	4
STAT 8330, Advanced Statistical Applications and Computing	Second Year, Fall	3
	Total	32

13) **Alternative Curricular Pathway:** What alternative curricular pathways exist (for example for students who were not admitted to the major but are still in satisfactory standing at the institutional level)? Please describe them below and describe how these students are advised about the alternative(s).

Students who are not admitted to the degree program will be advised about alternative complementary master's degree programs, such as in artificial intelligence, computer science, or statistics.

14) **Prior Learning Assessment:** Does the program include credit for prior learning assessment? How will credit be assessed and for what specific courses in the curriculum inclusive of prerequisites? If this is not applicable, indicate "NA" in this section.

N/A

15) **Open Educational Resources:** Does the program include open educational resources that have been assessed for quality and permissions, can be connected with related curricular resources, and are mapped to learning outcomes? If this is not applicable, indicate "NA" in this section.

N/A

16) **Waiver to Degree-Credit Hour** (if applicable):

N/A

17) Student Learning Outcomes: Student Learning outcomes and other associated outcomes of the proposed program (provide a narrative explanation).

All graduates earning the Data Science (M.S.) degree offered by the Franklin College of Arts and Sciences will learn the essential skills necessary to pursue careers in a variety of data-oriented companies [e.g., computing/internet companies (Google, Amazon, Facebook, IBM); engineering companies (Intel, Samsung, Boeing); finance/insurance companies (Goldman Sachs, AIG, Liberty Mutual); pharmaceutical companies (Johnson & Johnson)]; government/national labs (NASA, NIST, DoD); or pursue further graduate studies in Statistics, Computer Science, or other disciplines. All graduates will be able to:

- Develop and implement data analysis strategies based on sound principles of Statistics and Computer Science;
- Demonstrate and articulate appropriate statistical and computing strategies that can be used to extract evidence from data;
- Develop software and algorithms; design and manage a variety of databases and structures; process data in distributed environments;
- Collect and analyze the data using techniques from statistics, data mining, and machine learning;
- Provide visualizations of the data and build statistical models to facilitate inference and policy decisions; and
- Interpret results of statistical analysis and assist decision makers.

18) Assessment: Describe institutional programmatic assessments that will be completed to ensure academic quality, viability, and productivity.

All academic programs are reviewed annually to assess the program outcomes and student learning outcomes. Students completing Data Science (M.S.) are required to complete a capstone project, which requires applying combination of concepts from computer science and statistics to do advanced data analytics in a domain area. The capstone project course objectives will encompass the student learning outcomes for the program. In addition, the new degree will be assessed as part of the UGA comprehensive program review carried out every seven years.

19) Accreditation: Describe disciplinary accreditation requirements associated with the program (if applicable, otherwise indicate NA).

N/A

20) SACSCOC Institutional Accreditation: Is program implementation contingent upon SACSCOC action (e.g., substantive change, programmatic level change, etc.)? Please indicate Yes or No. ________No______________

ENROLLMENT SECTION *(Consult with Enrollment Management)*

21) Recruitment and Marketing Plan: What is the institution's recruitment and marketing plan? What is the proposed program's start-up timeline?

The departments will utilize a number of venues including New Dawgs, Orientation, Advising, and Majors Fair to recruit and advertise. To publicize the program, the departments will also use their websites, social media platforms such as Twitter and Facebook, and mail or email brochures and newsletters to potential feeder programs both at UGA and nationwide. This will begin in fall 2021 or as soon as USG approval is secured.

22) Enrollment Projections: Provide projected enrollments for the program specifically during the initial years of implementation.

a) Will enrollments be cohort-based? Yes____ or No__**X**___ (place an X beside one)

b) Explain the rationale used to determine enrollment projections.

As mentioned in Section 7, Demand, the pool of potential students for the proposed Data Science (M.S.) even within UGA is rather large since it would consist of students currently pursuing a B.S. in Computer Science (≈ 1,155) or Statistics (≈ 180) or Data Science (≈ 56), and some of those who wish to earn an M.S. degree in Computer Science (≈ 120) or Statistics (≈ 50). Of the total 90 students who responded to the survey referenced in Section 7, 44 indicated "Definitely Yes" or "Probably Yes" in pursuing an M.S. in Data Science. The Statistics department also expects a number of dual-degree students who wish to earn an M.S. in Statistics but are simultaneously earning an M.S. or a Ph.D. degree in another discipline within UGA to shift to the proposed Data Science (M.S.). Based on these and the survey results, the conservative enrollment projection for Year 1 (2021) of the new program is 15, which assumes that five of the existing M.S. students from Computer Science or Statistics will shift into the new program and 10 new students will enter the new program. Given that the Data Science (B.S.) major has doubled in enrollment since its introduction in fall 2019 (27 students in fall 2019 to about 56 students in fall 2020), the departments conservatively estimate new enrollments in the master's program at 15 new students for year 2, followed by a modest increase in the numbers during years 3 and 4, respectively, reaching a projected total number of 52 students in the program by 2024.

If projections are not met, the directors of the program—the heads of Computer Science and Statistics—along with the two respective graduate coordinators will develop a recruitment strategy to increase enrollment. The table below gives the enrollment projections for each year beginning fall 2021.

	Year 1 2021	Year 2 2022	Year 3 2023	Year 4 FY 2024

I. ENROLLMENT PROJECTIONS				
Student Majors				
Shifted from other programs	5	3	1	0
New to the institution	10	15	22	29
Total Majors	**15**	**33**	**41**	**52**

23) Faculty

a) Provide the total number of faculty members that will support this program: _45___
b) Submit your SACSCOC roster for the proposed degree. Annotate in parentheses the person who will have administrative responsibility for the program. Indicate whether any positions listed are projected new hires and currently vacant.

Faculty Name	Rank	Courses Taught (including term, course number & title, credit hours (D, UN, UT, G)	Academic Degrees & Coursework (relevant to courses taught, including institution & major; list specific graduate coursework, if needed)	Current Workload	Other Qualifications & Comments (related to courses taught)
T. N. Sriram	Professor (Department Head)	**Fall** STAT 4280/6280, Applied Time Series Analysis, 3.0 (UT/G)	Ph.D. Statistics, Michigan State University M.S. Statistics, University of Pune B.S. Statistics, Madras Christian College	3-6 credit hours/semester	**Ph.D. thesis:** Sequential Estimation of Parameters in a First Order Autoregressive Model

Jeongyuon Ahn	Associate Professor	**Spring and Fall** STAT 6510, Mathematical Statistics I, 3.0 (G) **Fall** STAT 6800, Tools for Statistical Theory, 3.0 (G) STAT 8210, Multivariate: Theory and Methods **Spring** STAT 8260, Theory of Linear Models, 4.0 (G)	Ph.D. Statistics, University of North Carolina, Chapel Hill M.S. Statistics, Seoul National University B.S. Statistics, Seoul National University	3-7 credit hours/semester	**Ph.D. thesis:** High Dimension, Low Sample Size Data Analysis
Gauri S. Datta	Professor	**Fall** STAT 4510, Mathematical Statistics I, 3.0 (G)	Ph.D. in Statistics, University of Florida M.Sc. in Statistics, University of Calcutta B.Sc. in Statistics, Ramakrishna Mission Residential College	3-6 credit hours/semester	**Ph.D. Thesis:** Bayesian prediction in mixed linear models: applications to small area estimation.
Mickey Dunlap	Lecturer	**Fall and Spring**	Ph.D., Statistics, Texas A&M	24 credit hours per academic	**Ph.D. Thesis**: Using the Bootstrap to

		STAT 2010, Statistical Methods for Data Scientists, 3.0 (UT) **Summer** STAT 4510/6510**,** Mathematical Statistics I**,** 3.0 (UT)	University, College Station, Texas M.S., Statistics, Mississippi State University, Mississippi State B.S., Mathematics, University of Tennessee at Martin	year	Analyze Variable Stars. **MS Thesis:** Comparison of Estimates of the Non- Centrality Parameter of a Non-Central Chi-Square.
Mohamad Kazem Shirani Faradonbeh	Assistant Professor	**Spring** STAT 4280/6280, Applied Time Series Analysis, 3.0 (UT/G)	Ph.D. in Statistics, University of Michigan M.A. in Statistics, University of Michigan B.S. in Electrical Engineering, Sharif University of Technology	3-6 credit hours/ semester	**Ph.D. Thesis** *Non-Asymptotic Adaptive Control of Linear-Quadratic Systems*

Daniel Hall	Professor	**Fall** STAT 8620, Categorical Data Analysis and Generalized Linear Models, 3.0 (UT/G) STAT 4620/6620, Applied Categorical Data Analysis, 3.0 (UT/G) **Spring** STAT 8630, Mixed-Effect Models and Longitudinal Data Analysis, 3.0 (UT/G)	Ph.D. in statistics, Northwestern University M.S. in statistics, Northwestern University B.A. in Mathematics, Northwestern University	3-6 credit hours/ semester	**Ph.D. Thesis** Extended Generalized Estimating Equations for Longitudinal Data
Pengsheng Ji	Associate Professor	**Spring** STAT 7000, Master's Research, 3.0 (UT/G)	Ph.D., Statistics, Cornell University M.S., Statistics, Cornell University M.S., Statistics and Probability, Nankai University B.S., Shiing-Shen Chern Special Class of Mathematics,	3-6 credit hours/sem ester	**Ph.D. Thesis:** Selected Topics in Nonparametric Testing and Variable Selection for High Dimensional Data

			Nankai University		
Yuan Ke	Assistant Professor	**Spring** STAT 4250/6250, Applied Multivariate Analysis and Statistical Learning, 3.0 (UT/G) **Fall** STAT 6510, Mathematical Statistics I, 3.0 (G)	Ph.D. in Mathematics – The University of York M.Sc. (Research) in Statistics – London School of Economics and Political Science B.Sc. in Physics, Beijing Normal University	3-6 credit hours/semester	**Ph.D. Thesis:** Feature Selection and Structure Specification in Ultra-High Dimensional Semi- Parametric Model with an Application in Medical Science
Liang Liu	Associate Professor	**Spring** STAT 8070, Statistical Computing II, 3.0 (G)	Ph.D. in Biostatistics, The Ohio State University M.S. in Statistics, The Ohio State University M.S. in Neuroscience, The Capital University of Medicine, Beijing, China B.S. in Clinical Medicine, Tianjin Medical University, Tianjin, China	3-6 credit hours/semester	**Ph.D. Thesis:** Reconstructing posterior distributions of a species phylogeny using estimated gene tree distributions
Ping Ma	Professor	**Spring** STAT 8070, Statistical	Ph.D. Statistics, Purdue University	3-6 credit hours/sem	**Ph.D. Thesis:** Nonparametric Mixed-effect Models

		Computing II, 3.0 (G) STAT 4280/6280, Applied Time Series Analysis, 3.0 (UT/G) STAT 8270, Spatial Statistics, 3.0 (G)	B.S. Mathematical Economics, Nankai University, China		Co-director of the Big Data Analytics Lab
Abhyuday Mandal	Professor	**Spring** STAT 6430, Design and Analysis of Experiments, 3.0 (G) **Fall** STAT 6420, Design and Analysis of Experiments, 3.0 (G)	Ph.D. Statistics, Georgia Institute of Technology M.A. Statistics, University of Michigan M.S. Statistics, Indian Statistical Institute B.S. Statistics, Indian Statistical Institute	3-6 credit hours/semes ter	**Ph.D. thesis:** Some Contributions to Design Theory and Applications
Cheolwoo Park	Professor	**Fall and Spring** STAT 4280/6280, Applied Time Series Analysis, 3.0 (UT/G) **Fall** STAT 8330, Advanced Statistical Applications	Ph.D. in Statistics, Seoul National University, South Korea MS in Statistics, Seoul National University " BS in Statistics, Seoul National University	3-6 credit hours/semes ter	**Ph.D. Thesis** Wavelet Estimation of Regression Functions with Sharp Change Points **M.S. Thesis:** On Kernel Estimates with Variable Bandwidths

		and Computing, 3.0 (G) STAT 7000, Master's Research, 3.0 (UT/G)			
Jaxk Reeves	Associate Professor	**Spring** STAT 8000, Introductory Statistics Collaboration, 3.0 (G) **Fall** STAT 6420, Design and Analysis of Experiments, 3.0 (G) STAT 8620, Categorical Data Analysis and Generalized Linear Models, 3.0 (G) **Summer** STAT 8000, Introductory Statistics Collaboration, 3.0 (G)	Ph.D. in Statistics, University of California, Berkeley B.S. in Applied Mathematics, Massachusetts Institute of Technology	3-6 credit hours/semester	**Ph.D. Thesis:** A Statistical Analysis and Projection of the Impact of Divorce on American Kinship Structures
Paul Schliekelman	Associate Professor	**Fall** STAT 8060, Statistical Computing I,	Ph.D. in Biomathematics, North Carolina State University Master of Biomathematics	3-6 credit hours/semester	PhD Thesis: Assessing New Methods for Autocidal Control

		3.0 (G) STAT 8330, Advanced Statistical Applications and Computing, 3.0 (G)	, North Carolina State University B.S. in Physics, Iowa State University		
Lynne Seymour	Associate Professor	**Fall** STAT 6510**,** Mathematical Statistics I**,** 3.0 (G) STAT 7000, Master's Research, 3.0 (UT/G) STAT 8060, Statistical Computing I, 3.0 (G) **Spring** STAT 8270, Spatial Statistics, 3.0 (G)	Ph.D. in Statistics. University of North Carolina, Chapel Hill. M.S. in Statistics. University of North Carolina M.S. in Mathematics. Auburn University B.S. in Mathematics. Troy State University.	3-6 credit hours/semester	**Ph.D. Thesis:** Parameter Estimation and Model Selection in Image Analysis using Gibbs-Markov Random Fields
Justin Strait	Assistant Professor	**Spring** STAT 4350/6350, Applied Bayesian Statistics, 3.0 (UT/G) **Spring and Fall** STAT 4510**,** Mathematical	Ph.D. in Statistics, The Ohio State University M.S. in Statistics, The Ohio State University B.S., Mathematics, Atmospheric Sciences,		**Ph.D. Thesis:** Elastic statistical shape analysis with landmark constraints

		Statistics I, 3.0 (UT)	University of Utah		
Mark Werner	Senior Lecturer	**Fall** STAT 4/6240, Sampling and Survey Methods, 3.0 (UT/G)	Ph.D. Applied Mathematics, University of Colorado Denver M.S. Applied Mathematics, University of Stellenbosch B.S. Honors, Applied Mathematics, University of Stellenbosch B.S. Applied Mathematics, University of Stellenbosch	18-24 credit hours per academic year	**Ph.D. Thesis:** Identification of Multivariate Outliers in Large Data Sets M.S. Thesis: Identifying Chaos in Experimental Data
Wenxuan Zhong	Professor	**Spring** STAT 8070, Statistical Computing II, 3.0 (G)	Ph.D. in Statistics, Purdue University B.S. Mathematical Economics, Nankai University, China	3-6 credit hours/semester	**Ph.D. Thesis** Nonparametric Clustering and Model Selection with Applications in Bioinformatics Co-direct the Big Data Analytics Lab; Founding Member of Georgia Informatics Institute

Thiab Taha	Professor (Department Head)	**Spring** CSCI 2150, Introduction to Computational Science, 4.0 (UT)	Ph.D. Applied Mathematics & Computer Science, Clarkson University M.Sc. Mathematics (Num. Analysis), University of Jordan B.Sc. Mathematics, University of Jordan	5 credit hours/year	Ph.D. dissertation: "On the Numerical and Analytical Aspects of Certain Nonlinear Evolution Equations"
Ismailcem Budak Arpinar	Associate Professor	**Spring** CSCI 4370/6370, Database Management, 4.0 (UT/G); CSCI 7200, Masters Project, 4.0 (G) **Fall** CSCI 4370/6370, Database Management, 4.0 (UT/G); CSCI 7200, Masters Project, 4.0 (G)	Ph.D. Computer Science, Middle East Technical University M.Sc. Computer Science, Middle East Technical University B.Sc. Computer Science, Middle East Technical University	5 credit hours/sem.	Ph.D. thesis: "Formalization of Workflows and Correctness Issues in Presence of Concurrency" M.Sc. thesis: "An Advanced Graphical User Interface for Object-Oriented DBMSs: Moodview"

Brad Barnes	Senior Lecturer (Undergrad. Coordinator)	**Spring** CSCI 1302, Software Development, 4.0 (UT) **Fall** CSCI 1302, Software Development, 4.0 (UT)	Ph.D. Computer Science, University of Georgia B.S. Computer Science, College of Charleston	8 credit hours/semes ter	Ph.D. dissertation: "A Regression Based System for Accurate Scalability Prediction on Large-Scale Machines"
Suchendra Bhandarkar	Professor	**Fall** CSCI 7200, Masters Project, 4.0 (G) **Spring** CSCI 8820 - Computer Vision and Pattern Recognition, 4.0 (G)	Ph.D., Syracuse University	5 credit hours/semes ter	

Liming Cai	Professor	**Spring** CSCI 4470/6470, Algorithms, 4.0 (UT/G) **Fall** CSCI 2670, Intro. Theory of Computing, 4.0 (UT) CSCI 7200, Masters Project, 4.0 (G)	Ph.D. Computer Science, Texas A&M University M.S. Computer Science, Tsinghua University, China B.S. Computer Science, Tsinghua University, China	5 credit hours/semester	Ph.D. dissertation: "Limited Nondeterminism and Approximation" M.S. thesis: "A Directed Translation System for Programming Languages"
Michael Cotterell	Lecturer	**Spring** CSCI 1302, Software Development, 4.0 (UT); CSCI 3030, Computing, Ethics, and Society, 3.0 (UT) **Fall** CSCI 1302, Software Development, 4.0 (UT)	Ph.D. Computer Science, University of Georgia B.S. Computer Science, University of Georgia	12 credit hours/semester	Ph.D. dissertation: "Supporting Open Science in Big Data Frameworks and Data Science Education."

Prashant Doshi	Professor	**Spring** CSCI 8920, Decision Making, 4.0 (G)	Ph.D., University of Illinois M.S., Drexel University B.E., University of Mumbai, India	5 credit hours/sem	Ph.D. dissertation: "Optimal Sequential Planning in Partially Observable Multiagent Settings"
Shelby Funk	Associate Professor	**Spring** CSCI 3030, Computing, Ethics, and Society, 3.0 (UT) **Fall** CSCI 3030, Computing, Ethics, and Society, 3.0 (UT); CSCI 4470/6470, Algorithms, 4.0 (UT/G)	Ph.D. Computer Science, University of North Carolina, Chapel Hill M.S. Computer Science, University of North Carolina, Chapel Hill M.S. Mathematics, University of North Carolina, Chapel Hill B.S. Mathematics, University of Maryland	5 credit hours/sem.	Ph.D. dissertation: "EDF Scheduling on Heterogeneous Multiprocessors" M.S. thesis: "Using a Modified Size Measure to Guide the Search in the Ordered Semantic Hyper-Linking Theorem Prover"

Le Guan	Assistant Professor	**Spring** CSCI 7200, Masters Project, 4.0 (G) **Fall** CSCI 7200, Masters Project, 4.0 (G)	Ph.D. Computer Science, Institute of Information Engineering, Chinese Academy of Sciences, China B.Eng. Computer Science and Engineering, University of Science and Technology of China, China	5 credit hours/sem.	
Bill Hollingsworth	Senior Lecturer	**Spring** CSCI 2150/L, Intro. Computational Science, 4.0 (UT) CSCI 2670, Intro. Theory of Computing, 4.0 (UT) **Fall** CSCI 2150, Intro. Scientific Computing, 4.0 (UT) CSCI 2670, Intro. Theory of Computing, 4.0 (UT)	Ph.D. Computer Science, University of Cambridge, U.K. M.Phil. Theoretical Linguistics, University of Cambridge, U.K. M.A. Mathematics, University of Georgia B.S. Mathematics, University of Georgia	12 credit hours/sem.	Ph.D. dissertation: “Using Lexical Chains to Characterize Scientific Text” M.Phil. thesis: “Rhythmic and Segmental Contributions to Timing in Speech Synthesis”

Maria Hybinette	Associate Professor	**Spring** CSCI 7200, Masters Project, 4.0 (G) **Summer** CSCI 7200, Masters Project, 4.0 (G) **Fall** CSCI 4210/6210, Simulation and Modeling, 4.0 (UT/G)	Ph.D. Computer Science, Georgia Institute of Technology M.S., Computer Science, Georgia Institute of Technology B.S. Mathematics and Computer Science, Emory University	5 credit hours/sem.	Ph.D. dissertation: "Interactive Parallel Simulation Environments"
In Kee Kim	Assistant Professor	**Spring** CSCI 4795/6795, Cloud Computing, 4.0 (UT/G) CSCI 7200, Masters Project, 4.0 (G)	Ph.D. Computer Science, University of Virginia M.S. Computer Science and Engineering, Inha University, South Korea B.S. Computer Science and Engineering, Inha University, South Korea	5 credit hours/sem.	Ph.D. dissertation: "Proactive Resource Provisioning to Ensure Predictable End-to-End Performance for Cloud Applications."

Krzysztof Kochut	Professor	**Spring** CSCI 7200, Masters Project, 4.0 (G) **Summer** CSCI 1302, Software Development, 4.0 (U) **Fall** CSCI 7200, Masters Project, 4.0 (G)	Ph.D. Computer Science, Louisiana State University M.S. Computer Science, University of Warsaw, Poland	5 credit hours/sem.	Ph.D. dissertation: "Resolution Proof Technique in Linear Temporal Logic" M.S. thesis: "Augmented Transition Network Grammar for Parsing of the Polish Language"
Salvatore Lamarca	Limited-Term Lecturer	**Spring** CSCI 1301, Intro. Computing and Programming, 4.0 (UT) **Summer** CSCI 1301, Intro. Computing and Programming, 4.0 (UT) **Fall** CSCI 1301, Intro. Computing and Programming, 4.0 (UT)	Ph.D. Computer Science, University of Georgia (IN PROGRESS) B.S. Computer Science, University of Georgia B.S. Mathematics, University of Georgia	12 credit hours/sem.	

Jaewoo Lee	Assistant Professor	**Spring** CSCI 8960, Privacy-Preserving Data Analysis, 4.0 (G)	Ph.D. Computer Science, Purdue University M.S. Computer Science, Yonsei University, South Korea	5 credit hours/sem.	Ph.D. dissertation: "Achieving Practical Differential Privacy" M.S. thesis: "Efficiently Tracing Clusters over High-dimensional Data Streams"
Kyu Hyung Lee	Associate Professor	**Spring** CSCI 7200, Masters Project, 4.0 (G) **Fall** CSCI 7200, Masters Project, 4.0 (G)	Ph.D. Computer Science, Purdue University M.S. Computer Engineering, Hong-Ik University, South Korea B.S. Computer Engineering, Hong-Ik University, South Korea	5 credit hours/sem.	Ph.D. dissertation: "Selective Logging for Accurate, Space Efficient Forensic Analysis and Reducible Execution Replay" M.S. thesis: "PFC: Transparent Optimization of Existing Prefetching Strategies for Multi-level Storage Systems"

Sheng Li	Assistant Professor	**Spring** CSCI 8950, Machine Learning, 4.0 (G) **Fall** CSCI 7200, Masters Project, 4.0 (G) CSCI 8945, Adv. Representation Learning, 4.0 (G)	Ph.D. Computer Engineering, Northeaster University M.Eng. Information Security, Nanjing University of Posts and Telecommunica tions, China B.Eng. Computer Science and Engineering, Nanjing University of Posts and Telecommunica tions, China	5 credit hours/sem.	Ph.D. dissertation: “Robust Data Representations for Visual Learning” M.Eng. thesis: “Divide and Conquer based Discrimination Feature Extraction and Its Application” B.Eng. thesis: “Kernel DCT Discriminant Analysis for Face Recognition”
Tianming Liu	Distinguished Research Professor	**Spring** CSCI 7200, Masters Project, 4.0 (G) **Fall** CSCI 7200, Masters Project, 4.0 (G) CSCI 8850, Adv. Biomedical Image Analysis, 4.0 (G)	Ph.D. Computer Science, Shanghai Jiaotong University, China M.S. Automation, Northwestern Polytechnical University, China B.S., Automation, Northwestern Polytechnical University, China	5 credit hours/sem.	Ph.D. dissertation: “On Adaptive Rate Control for Video Streaming” M.S. thesis: “Artificial Neural Networks for Object Detection”

John Miller	Professor (Graduate Coordinator)	**Spring** CSCI 4360/6360, Data Science II, 4.0 (UT/G); CSCI 7200, Masters Project, 4.0 (G) **Summer** CSCI 4370/6370, Database Management, 4.0 (UT/G) **Fall** CSCI 8370, Adv. Database Systems, 4.0 (G)	Ph.D. Information and Computer Science, Georgia Institute of Technology M.S. Information and Computer Science, Georgia Institute of Technology B.S. Applied Mathematics, Northwestern University	5 credit hours/sem.	Ph.D. dissertation: "Dissertation was in Markovian Analysis and Optimization of Database Recovery Protocols"

Ramviyas Nattanmai Parasuraman	Assistant Professor	**Spring** CSCI 8535, Multi-Robot Systems, 4.0 (G) CSCI 7200, Masters Project, 4.0 (G)	Ph.D. Robotics and Automation, Universidad Politécnica de Madrid (Technical University of Madrid), Madrid, Spain M.Tech. Instrument Technology, Indian Institute of Technology Delhi, New Delhi, India B.Eng. Electronics and Instrumentation, Thiagarajar College of Engineering (Anna University), Madurai, India	5 credit hours/sem.	
Hao Peng	Lecturer	**Spring** CSCI 1301, Intro. Computing and Programming, 4.0 (UT) CSCI 3360, Data Science I, 4.0 (UT)	Ph.D. Computer Science, University of Georgia B.S. Computer Science, University of Georgia B.S. Statistics, University of Georgia	12 credit hours/sem.	

Shannon Quinn	Assistant Professor	**Fall** CSCI 4360/6360, Data Science II, 4.0 (UT/G) CSCI 4360/6360, Data Science II, 4.0 (UT/G)	Ph.D. Computational Biology, University of Pittsburgh M.S. Computational Biology, Carnegie Mellon University B.S. Computer Science, Georgia Institute of Technology	5 credit hours/sem.	Ph.D. dissertation: "Distributed Spectral Graph Methods for Analyzing Large-Scale Unstructured Biomedical Data" M.S. thesis: “Waldo: A Framework for Inferring Protein Location as a Function of Condition"
Lakshmish Ramaswamy	Professor	**Spring** CSCI 4780/6780, Distributed Computing Systems, 4.0 (UT/G) CSCI 7200, Masters Project, 4.0 (G) **Fall** CSCI 8790, Adv. Topics in Data Intensive Computing, 4.0 (G)	Ph.D. Computer Science, Georgia Institute of Technology M.S. Computer Science and Automation, Indian Institute of Science, India B.E. Computer Science and Engineering, University of Mysore, India	5 credit hours/sem.	Ph.D. thesis: “Towards Efficient Delivery of Dynamic Web Content” M.S. thesis: “Wavelets for Volume Graphics”

Khaled Rasheed	Professor	**Spring** CSCI (ARTI) 8950, Machine Learning, 4.0 (G) CSCI 7200, Masters Project, 4.0 (G) **Fall** CSCI 4380/6380, Data Mining, 4.0 (UT/G)	Ph.D. Computer Science, Rutgers University M.S. Computer Science, Rutgers University B.S. Computer Science, Alexandria University, Egypt	5 credit hours/sem.	Ph.D. dissertation: "GADO: A Genetic Algorithm for Continuous Design Optimization"
Eman Saleh	Lecturer	**Fall** CSCI 1302, Software Development, 4.0 (UT) CSCI 4050/6050, Software Engineering, 4.0 (UT/G)	Ph.D. Computer Science, University of Cairo, Egypt M.S. Computer Science, University of Jordan, Jordan B.S. Computer Science, University of Jordan, Jordan	12 credit hours/sem.	Ph.D. dissertation: “Model-Driven Engineering Design Approach for Developing Multi-Platform User Interfaces”

Michael Scott	Lecturer	**Spring** CSCI 1301/L, Intro. Computing and Programming, 4.0 (UT)	Ph.D. Computer Science, University of Georgia	12 credit hours/sem.	
Wenwen Wang	Assistant Professor	**Fall** CSCI 7200, Masters Project, 4.0 (G)	Ph.D. Computer Science, University of Chinese Academy of Sciences	5 credit hours/sem.	

c) Does the institution require additional faculty to establish and implement the program? Yes or No. _______No_______________ Please indicate your answer in the space provided.

Describe the institutional plan for recruiting additional faculty members in terms of required qualifications, financial preparations, timetable for adding faculty, and whether resources were shifted from other academic units, programs, or derived from other sources. Explain clearly whether additional faculty hires can be supported with institutional funds.

The Department of Statistics hired two top quality faculty in 2013, both working in Big Data and generating over 3 million in extramural funding from the National Institute of Health and National Science Foundation. Both of them are members of the new Georgia Informatics Institutes for Research and Education. In addition, they also maintain a Big Data Analytics Lab that supports about 10 graduate students each year. Statistics has also hired a tenure-track faculty member and a non-tenure-track faculty member each year since 2015. In addition, from the 2018 and 2019 Presidential Hiring Initiatives, Statistics has hired two tenure-track faculty in statistical machine learning and data science, respectively, and a lecturer in the area of statistics and business statistics. Similarly, The Computer Science department has added 11 top quality faculty since 2015. With these new positions in Computer Science and Statistics, this program is designed to be self-supportive at this time. In subsequent years, new faculty positions will be justified by increased enrollments.

24) Fiscal, Tuition, and Estimated Budget

a) Describe the resources that will be used specifically for the program.

All resources needed for the program are pre-existing. The departments will utilize the current resources (personnel, library, equipment, laboratory, and computing) available at the departmental and university level.

b) Does the program require a tuition cost structure different from or above a regular tuition designation for the degree level? Yes _______ or No ___**X**____ (place an X beside one)

c) Does the program require a special fee for the proposed program? Yes ____ or No__**X**___ (place an X beside one)

d) If the program requires a different tuition cost structure or special fee, such requests require approval through both the Committee on Academic Affairs (for the academic program) and the Committee on Fiscal Affairs (for the tuition increase or special fee designation). The resultant tuition and/or fee request for a new degree is to be submitted to both the academic affairs and fiscal affairs offices. Complete Appendix III that includes information for a differential tuition cost structure involving a proposal for a new academic program.

N/A

e) Note: The web link for approved tuition and fees for USG institutions is located at the following url: http://www.usg.edu/fiscal_affairs/tuition_and_fees

f) Budget Instructions: Complete the form further below and **provide a narrative to address each of the following**:

g) For Expenditures (*ensure that the narrative matches the table*):

i. Provide a description of institutional resources that will be required for the program (e.g., personnel, library, equipment, laboratories, supplies, and capital expenditures at program start-up and recurring).

- All faculty resources needed for the program are pre-existing. Given that this program will be jointly managed by the Department of Computer Science and Department of Statistics, the faculty members from both the departments will fulfill the instructional needs in the proposed program. The current infrastructure will be used, and staff support provided to both departments.
- Personnel expenditures for each fiscal year are calculated using average per course instructional cost associated with offering major required courses each year. The average instructional cost for the new course in the major is taken to be $18,000, whereas for courses that are common to the Data Science (DS) major and the Statistics and Computer Science (CS) majors, the average instructional cost is $5,000. The average instructional cost is calculated using the average faculty salary multiplied by the average instructional EFT and divided by the average course load.

- For the first year, the expenditure is determined based on offering three required courses (CSCI(STAT) 6375 (NEW), CSCI 6360, STAT 6420) and three CSCI or STAT elective courses (e.g., CSCI 6370, STAT 6250, and STAT 8000). In the second year, the expenditure is based on offering one required course (STAT 8330), one CSCI or STAT elective course (CSCI 8360 or STAT 6280), and a required Master's Project (CSCI 7200 or STAT 7000).

ii. If the program involves reassigning existing faculty and/or staff, include the specific costs/expenses associated with reassigning faculty and staff to support the program (e.g., cost of part-time faculty to cover courses currently being taught by faculty being reassigned to the new program, or portion of full-time faculty workload and salary allocated to the program).

The departments will be shifting teaching assignment for one of the existing faculty members from teaching a lower division CSCI or STAT course to a Graduate Teaching Assistant (GTA) for the first few years of the program. The two participating departments will cover the cost of the GTA. As enrollments increase, new positions are anticipated for this proposed program.

h) For Revenue (*ensure that the narrative matches the table*):

i. If using existing funds, provide a specific and detailed plan indicating the following three items: source of existing funds being reallocated; how the existing resources will be reallocated to specific costs for the new program; and the impact the redirection will have on units that lose funding.

- **Source of existing funds being reallocated**
  Existing faculty lines budgeted for instruction will be utilized to cover instructional costs associated with courses that are common to this degree and the Statistics and Computer Sciences degrees. Funds to teach the one new course for the major will be reallocated using the new faculty lines which are already allocated to both departments.

- **How the existing resources will be reallocated to specific costs for the new program**
  Instructional time for existing and newly authorized faculty lines will be used to cover the needed program instruction.

- **The impact the redirection will have on units that lose funding**
  No funding or instruction will be lost as a result of this program.

ii. Explain how the new tuition amounts are calculated.

Tuition is calculated based on the 2020-2021 University of Georgia rate for Master's students of $370/credit hour or a rate of $4,439 for 12 or more credit hours/semester. For each year, the anticipated number of students enrolled in the program during the fall and spring, respectively, is multiplied by $4,439 to obtain the total tuition per year. For example, it is anticipated that 15 students will be enrolled in the program during the first year. Therefore, for Year 1, following Section 12, Program of Study, the total tuition amount is 15 students x 10 credit hours x $370 (Fall) + 15 students x 11 credit hours x $370 (Spring) = $116,550. Similarly, the tuition amounts for years 2 to 4 were calculated using the respective anticipated enrollment numbers (see calculations in the table below).

iii. Explain the nature of any student fees listed (course fees, lab fees, program fees, etc.). Exclude student mandatory fees (i.e., activity, health, athletic, etc.).

There are no new student or program fees.

iv. If revenues from Other Grants are included, please identify each grant and indicate if it has been awarded.

N/A

v. If Other Revenue is included, identify the source(s) of this revenue and the amount of each source.

N/A

i) Revenue Calculation: Provide the revenue calculation, in other words, the actual calculation used to determine the projected tuition revenue amounts for each fiscal year involving start-up and implementation of the proposed program.

Year One
Year One students – (Fall: 15 students x 10 credit hours x $370) + (Spring: 15 students x 11 credit hours x $370) = $55,500 + $61,050 = $116,550

	Year One	**Year Two**	**Year Three**	**Year Four**
Year One 15 students	$116,550	$61,050	N/A	N/A
Year Two 18 students	N/A	$139,860	$73,260	N/A
Year Three 23 students	N/A	N/A	$178,710	$93,610
Year Four 29 students	N/A	N/A	N/A	$225,330
Total	$116,550	$200,910	$251,970	$318,940

Year Two
Year One students – (Fall: 15 students x 11 credit hours x $370) = $61,050
Year Two students – (Fall: 18 students x 10 credit hours x $370) + (Spring: 18 students x 11 credit hours x $370) = $66,600 + $73,260 = $139,860

Year Three
Year One students – graduated
Year Two students – (Fall: 18 students x 11 credit hours x $370) = $73,260
Year Three students – (Fall: 23 students x 10 credit hours x $370) + (Spring: 23 students x 11 credit hours x $370) = $85,100 + $93,610 = $178,710

Year Four
Year One students – graduated
Year Two students – graduated
Year Three students – (Fall: 23 students x 11 credit hours x $370) = $93,610
Year Four students – (Fall: 29 students x 10 credit hours x $370) + (Spring: 29 students x 11 credit hours x $370) = $107,300 + $118,030 = $225,330

j) When Grand Total Revenue is not equal to Grand Total Costs:
   i. Explain how the institution will make up the shortfall. If reallocated funds are the primary tools being used to cover deficits, what is the plan to reduce the need for the program to rely on these funds to sustain the program?

      N/A – There is no shortfall because there is no new cost as a result of this program.

   ii. If the projected enrollment is not realized, provide an explanation for how the institution will cover the shortfall.

      If the enrollments do not match projections, there will be no budget shortfall and there would be no additional cost to the University because the existing courses will continue to be taught for the Statistics and Computer Science majors, and the new courses for the Data Science major will not be offered.

   i. If the projected enrollment is not realized, what are your next action steps in terms of bolstering the program, potentially altering the program, teach-outs, a planned phase-out, etc.?

      If the projected enrollment for this program is not realized, the departments will increase advertising for the program, as well as survey students who have enrolled in the major for recommendations on how to attract additional students.

I. EXPENDITURES	First FY Dollars	Second FY Dollars	Third FY Dollars	Fourth FY Dollars
Personnel – reassigned or existing positions				
Faculty **(see 23.g.ii)**	359,088	369,860	380,956	392,385
Part-time Faculty **(see 23.g.ii)**				
Graduate Assistants **(see 23.g.ii)**				
Administrators **(see 23.g.ii)**				
Support Staff **(see 23.g.ii)**				
Fringe Benefits	118,499	122,054	125,716	129,487
Other Personnel Costs				
Total Existing Personnel Costs	477,587	491,914	506,672	521,872

EXPENDITURES (Continued)				
Personnel – new positions (see 23.g.i)				
Faculty		20,471	22,716	43,457
Part-time Faculty				
Graduate Assistants	20,031	20,632	21,251	
Administrators				
Support Staff				
Fringe Benefits		6,844	7,496	14,341
Other personnel costs				
Total New Personnel Costs	**20,031**	**48,217**	**51,463**	**57,798**
Start-up Costs (one-time expenses) (see 23.g.i)				
Library/learning resources				
Equipment				
Other				

Physical Facilities: construction or renovation (see section on Facilities)				
Total One-time Costs	**0**	**0**	**0**	**0**
Operating Costs (recurring costs – base budget) (see 23.g.i)				
Supplies/Expenses				
Travel				
Equipment				
Library/learning resources				
Other				
Total Recurring Costs				
GRAND TOTAL COSTS	497,618	540,131	558,135	579,669

III. REVENUE SOURCES				
Source of Funds				
Reallocation of existing funds **(see 23.h.i)**	497,618	491,914	506,672	521,872
New student workload				
New Tuition **(see 23.h.ii)**	$116,550	$200,910	$251,970	$318,940
Federal funds				
Other grants **(see 23.h.iv)**				
Student fees **(see 23.h.iii)** Exclude mandatory fees (i.e., activity, health, athletic, etc.).				
Other **(see 23.h.v)**				

New state allocation requested for budget hearing				
GRAND TOTAL REVENUES	$594,137	692,824	758,642	840,812
Nature of Revenues				
Recurring/Permanent Funds				
One-time funds				
Projected Surplus/Deficit (Grand Total Revenue – Grand Total Costs) (see 20.h.i. & 20.h.ii).	96,519	152,693	200,507	261,142

25) Facilities/Space Utilization for New Academic Program Information

Facilities Information — Please Complete the table below.

				Total GSF
a.	**Indicate the floor area required for the program in gross square feet (gsf). When addressing space needs, please take into account the projected enrollment growth in the program over the next 10 years.**			10,000
b.	**Indicate if the new program will require new space or use existing space. (Place an "x" beside the appropriate selection.)**			
	Type of Space		**Comments**	
i.	Construction of new space is required (x).- □	N/A		
ii.	Existing space will require modification (x). □	N/A		
iii.	If new construction or renovation of existing space is anticipated, provide the justification for the need.		N/A	
iv.	Are there any accreditation standards or guidelines that will impact facilities/space needs in the future? If so, please describe the projected impact.		No	
v.	Will this program cause any impact on the campus infrastructure, such as parking, power, HVAC, other? If yes, indicate the nature of the impact, estimated cost, and source of funding.		No	
vi.	Indicate whether existing space will be used.	X	Existing facilities will be sufficient	
c.	**If new space is anticipated, provide information in the spaces below for each category listed:**			
i.	Provide the estimated construction cost.		N/A	
ii.	Provide the estimated total project budget cost.		N/A	
iii.	Specify the proposed funding source.		N/A	

iv.	What is the availability of funds?	N/A
v.	When will the construction be completed and ready for occupancy? (Indicate semester and year).	N/A
vi.	How will the construction be funded for the new space/facility?	N/A
vii.	Indicate the status of the Project Concept Proposal submitted for consideration of project authorization to the Office of Facilities at the BOR. Has the project been authorized by the BOR or appropriate approving authority?	N/A

d.	**If existing space will be used, provide information in the space below.**
	Provide the building name(s) and floor(s) that will house or support the program. Indicate the campus, if this is part of a multi-campus institution and not physically located on the main campus. Please do not simply list all possible space that could be used for the program. We are interested in the actual space that will be used for the program and its availability for use.
	Boyd Graduate Studies building (home of Computer Science) and Brooks Hall (4th and 5th Floor; home of Statistics) will house and support the program. Classroom spaces on south campus will be used for computer science classes, whereas classrooms in Caldwell Hall, Sanford Hall, and Park Hall will be used for Statistics classes.

e. List the specific type(s) and number of spaces that will be utilized (e.g. classrooms, labs, offices, etc.)

i.	**No. of Spaces**	**Type of Space**	**Number of Seats**	**Assignable Square Feet (ASF)**
	5	Classrooms	40	12,500
	2	Labs (dry)	50	6,250
		Labs (wet)		
		Meeting/Seminar Rooms		
		Offices		

		Other (specify)		
Total Assignable Square Feet (ASF)				**18,750**
ii.	If the program will be housed at a temporary location, please provide the information above for both the temporary space and the permanent space. Include a time frame for having the program in its permanent location.			

Chief Business Officer or Chief Facilities Officer Name & Title	**Phone No.**	**Email Address**
	Signature	
Note: A Program Manager from the Office of Facilities at the System Office may contact you with further questions separate from the review of the new academic program.		

FINAL NOTE:

Appendices that do not apply to the proposed program should not be attached.

APPENDIX I

Use this section to include letters of support, curriculum course descriptions, and recent rulings by accrediting bodies attesting to degree level changes for specific disciplines, and other information.

Appendix A: Sample Program of Study

Courses (list acronym, number, and title)	**Semester**	**Hours**
CSCI(STAT) 6375, Foundations of Data Science (NEW)	First Year, Fall	4
STAT 6420, Applied Linear Models	First Year, Fall	3
STAT 6250, Applied Multivariate Analysis and Statistical Learning	First Year, Fall	3
CSCI 6360, Data Science II	First Year, Spring	4
CSCI 6370, Database Management	First Year, Spring	4
STAT 8000, Introductory Statistical Collaboration	First Year, Spring	3
CSCI 8360, Data Science Practicum	Second Year, Fall	4
CSCI 7200, Masters Project	Second Year, Fall	4
STAT 8330, Advanced Statistical Applications and Computing	Second Year, Fall	3
	Total	32

Appendix II: Course Titles, Credits, Descriptions

CSCI 6150
Credits: 4
Course Title: Numerical Simulations in Science and Engineering
Course Description: Computationally oriented, covering a wide range of topics that are necessary for numerical simulation in science and engineering. Sequential and parallel numerical methods will be introduced. Available symbolic and numerical software packages (e.g., Matlab, Maple and MPI) and visualization tools will be used in the mathematical simulations.
Prerequisites: (MATH 2250 and CSCI 1301-1301L) or permission of department

CSCI 6170
Credits: 4
Course Title: Introduction to Computational Investing
Course Description: An introduction to implementing computational-based trading strategies from information gathering to market ordering and trading, including probabilistic machine-learning approaches to situational analysis and to trading decisions. We consider approaches like linear regression, decision trees, K nearest neighbors, and reinforcement learning and apply them to real- world trading.
Prerequisite: CSCI 2720

CSCI 6210
Credits: 4
Course Title: Simulation and Modeling
Course Description: The modeling and simulation of existing or planned systems for the purpose of studying their correctness, reliability, or performance. Topics to be addressed include discrete-event simulation, continuous simulation, analysis and modeling methodologies, animation, virtual reality, and Web-based simulation.
Prerequisite: CSCI 2720

CSCI 6260
Credits: 4
Course Title: Data Security and Privacy
Course Description: Examination of security and privacy issues related to protecting personal data in various environments (for example: cloud computing, smart grid, and internet of things) cover the fundamentals and principles of data security and privacy, and computational and statistical techniques for constructing secure and private systems.
Prerequisite: CSCI 1302

CSCI (STAT) 6375
Credits: 4
Course Title: Foundations of Data Science (NEW)
Course Description: Introduction to data life cycle including collection, transmission, storage, retrieval, provenance visualization and analysis. A rigorous overview of data

quality, indexing and querying techniques, data mining and scientific computing methodologies and associated computing infrastructures. Core concepts in supervised and unsupervised analytics, dimensionality reduction, and data visualization will be explored.
Prerequisite: None

CSCI 6360
Credits: 4
Course Title: Data Science II
Course Description: An introduction to advanced analytics techniques in Data Science, including random forests, semi-supervised learning, spectral analytics, randomized algorithms, and just-in-time compilers. Distributed and out-of-core processing
Prerequisite: CSCI 3360

CSCI 6370
Credits: 4
Course Title: Database Management
Course Description: The theory and practice of database management. Topics to be covered include efficient file access techniques, the relational data model as well as other data models, query languages, database design using entity-relationship diagrams and normalization theory, query optimization, and transaction processing.
Prerequisite: CSCI 2720

CSCI 6380
Credits: 4
Course Title: Data Mining
Course Description: A broad introduction to data mining methods and an exploration of research problems in data mining and its applications in complex real-world domains. Approaches include association and classification rule learning, tree learning, neural network and Bayesian methods, support vector machines, clustering, and ensemble learning.
Prerequisite: CSCI 2720

CSCI 6470
Credits: 4
Course Title: Algorithms
Course Description: Algorithms, covering basic analysis techniques, basic design techniques (divide-and-conquer, dynamic programming, greedy, and branch-and-bound), basic graph algorithms, and NP-completeness.
Prerequisite: CSCI 2720

CSCI (MATH) 6690
Credits: 3
Course Title: Graph Theory
Course Description: Elementary theory of graphs and digraphs. Topics include connectivity, reconstruction, trees, Euler's problem, hamiltonicity, network flows,

planarity, node and edge colorings, tournaments, matchings, and extremal graphs. A number of algorithms and applications are included.
Prerequisite: CSCI(MATH) 2610 or MATH 3200

CSCI 6780
Credits: 4 hours
Course Title: Distributed Computing Systems
Course Description: The fundamental concepts in distributed computing and the practical techniques for building distributed systems. Topics include distributed computing models, naming, synchronization, replication and consistency, fault tolerance, and security. Widely deployed distributed systems are used as case studies. Students design, implement, and analyze prototype systems.
Prerequisite: CSCI 2720 and CSCI 1730

CSCI 6795
Credits: 4 hours
Course Title: Cloud Computing
Course Description: Introduction to cloud computing for undergraduate and graduate students. Topics include virtualization, data centers, virtual machines, cloud service models, public vs. private vs. hybrid clouds, open stack, container and orchestration infrastructure, cloud storage, mobile cloud, IoT, and big data. Students also learn how to design, implement, and test cloud applications.
Prerequisite: CSCI 2720

CSCI 6850
Credits: 4
Course Title: Biomedical Image Analysis
Course Description: Introduction to the standard approaches to biomedical image analysis, including basic concepts of biomedical imaging, basic algorithms, principles of software systems, and their applications. Biomedical image analysis software tools will be used in hands-on projects.
Prerequisite: CSCI 2720 or permission of department

CSCI 7007
Credits: 3 hours. Repeatable for maximum 12 hours credit.
Course Title: Internship in Computer Science Business/Industry
Course Description: Internship in a professional setting allowing the student to integrate his/her educational experience with real-world situations in software development, database technology, hardware design, networks, etc.
Prerequisite: Permission of department

CSCI 7200
Credits: 2-6 hours. Repeatable for maximum 12 hours credit.
Course Title: Master's Project
Course Description: Applied research project under the direction of the major professor for the Computer Science Non-thesis MS degree. As part of the requirements, a

comprehensive report must be prepared detailing the student's procedures and findings regarding the completed project work.
Prerequisite: Permission of department

CSCI 8360
Credits: 4 hours
Course Title: Data Science Practicum
Course Description: Advanced data science techniques for analyzing large-scale data in distributed environments. Students will develop scalable algorithms in frameworks such as Spark and Flink. This course is team-based, involving several mini-projects over the course of the semester with a competition as the final project.
Prerequisite: CSCI 4360/6360 or CSCI 4380/6380 or permission of department

CSCI 8370
Credits: 4 hours
Course Title: Advanced Database Systems
Course Description: Advanced study of database systems. The course focuses on concepts, algorithms and technologies for relational, object-oriented and distributed database systems. Related technologies such as data warehouses and repositories will also be covered.
Prerequisite: CSCI 4370/6370 and CSCI 4730/6730

CSCI 8380
Credits: 4 hours
Course Title: Advanced Topics in Information Systems
Course Description: Advanced topics in information systems and databases. The two major issues dealt with are: (1) information integration and interoperability, and (2) novel database technologies. The first addresses the integration of autonomous and heterogeneous resources managing structured, semi-structured, and unstructured data. The second deals with the query formulation, and processing on heterogeneous content. Special attention will be given to emerging research areas fueled by the Web and related technologies.
Prerequisite: CSCI 4370/6370 or CSCI 4050/6050

CSCI 8535
Credits: 4 hours
Course Title: Multi-Robot Systems
Course Description: This is primarily a research oriented, seminar-style course covering the topics of control, communication, cooperation, and coordination aspects in multi-robot systems. It enables students to understand, devise, and solve problems in multi-robot systems and will include project-based assignments.
Prerequisites: CSCI (ARTI) 4530/6530 and (CSCI 4500/6500 or CSCI 1730 or Permission of Department)

CSCI 8790
Credits: 4 hours
Course Title: Advanced Topics in Data Intensive Computing
Course Description: Modern computing applications require storage, management, and processing of petabytes of data. The data is not only extremely diverse, ranging from unstructured text and relational tables to complex graphs, but it is also dynamic. This course focuses on developing scalable architectures, algorithms, and techniques for supporting various data intensive applications.
Prerequisite: CSCI 4370/6370 or permission of department

CSCI 8820

Credit: 4 hours
Course Title: Computer Vision and Pattern Recognition
Course Description: Low-level and high-level vision including edge detection, connected component labeling, boundary detection, segmentation, stereopsis, motion analysis, and object recognition. Knowledge representation, knowledge retrieval and reasoning techniques in computer vision. Parallel computing, parallel architectures and neural computing for computer vision.
Prerequisites: CSCI 4810/6810 or Permission of Department

CSCI 8850
Credits: 4 hours
Course Title: Advanced Biomedical Image Analysis
Course Description: Introduction to advanced approaches to, and applications of, biomedical image analysis, including imaging biomarker discovery, computer-aided diagnosis, computer-aided follow-up, image-guided therapy, molecular imaging, functional imaging, and translational imaging. Advanced applications of state-of- the-art biomedical image analysis software systems will be emphasized.
Prerequisite: CSCI 4850/6850 or permission of department

CSCI 8920
Credits: 4 hours
Course Title: Decision Making Under Uncertainty
Course Description: Choosing optimally among different lines of actions is a key aspect of autonomy in artificial agents. This course will focus on how to make optimal and approximately optimal decisions in single and multiagent settings. It will be self-contained, introducing background literature such as aspects of probability and game theories.
Prerequisite: CSCI 4470/6470 or permission of department.

CSCI 8945
Credits: 4 hours
Course Title: Advanced Representation Learning
Course Description: Advanced representation learning algorithms in machine learning, from the traditional subspace learning models to the recent deep representation learning

models. Applications in the fields of computer vision, data mining, and natural language processing will be covered.
Prerequisite: (CSCI 4380/6380 and CSCI(PHIL) 4550/6550) or permission of department

CSCI 8950
Credits: 4 hours
Course Title: Machine Learning
Course Description: An in-depth introduction to machine learning methods and an exploration of research problems in machine learning and its applications which may lead to work on a project or a dissertation.
Prerequisite: CSCI(PHIL) 4550/6550 or CSCI 4560/6560 or permission of department

CSCI 8951
Credits: 4 hours
Course Title: Large-Scale Optimization for Machine Learning
Course Description: Introduction of optimization algorithms suitable for solving large-scale problems, with a focus on exploring recent advances in the context of machine learning. Students will learn several algorithms for solving smooth and non-smooth problems, compare the efficacy of those methods, and discuss the trade-offs in terms of statistical accuracy, scalability, and algorithmic complexity
Prerequisite: CSCI 4380/6380 or CSCI 4360/6360 or permission of department

CSCI 8955
Credits: 4 hours
Course Title: Advanced Data Analytics: Statistical Learning and Optimization
Course Description: Advanced topics in data analysis, with an emphasis on statistical learning and related optimization problems. The applications include regression, classification, and other tasks in image analysis. The lectures will be based on books and articles in the field of computer vision and medical image analysis.
Prerequisite: CSCI 4150/6150 or CSCI 4380/6380 or permission of department.

CSCI 8960
Credits: 4 hours
Course Title: Privacy-Preserving Data Analysis
Course Description: An introduction to the privacy preservation problems, as well as algorithmic and statistical techniques for data privacy, in modern data analysis, such as machine learning and data mining. Approaches include randomized algorithms, synthetic data generation, stability analysis, and so on.
Prerequisite: CSCI 4380/6380 or permission of department

STAT 6240
Credits: 3
Course Title: Sampling and Survey Methods
Course Description: Design of finite population sample surveys. Stratified, systematic, and multistage cluster sampling designs. Sampling with probability proportional to size. Auxiliary variables, ratio and regression estimators, non-response bias.

Prerequisites: STAT 6220 or STAT 6310 or STAT 6315 or permission of department

STAT 6250
Credits: 3
Course Title: Applied Multivariate Analysis and Statistical Learning
Course Description: The methodology of multivariate statistics and machine learning for students specializing in statistics. Topics include inference on multivariate means, multivariate analysis of variance, principal component analysis, linear discriminant analysis, factor analysis, linear discrimination, classification trees, multi-dimensional scaling, canonical correlation analysis, clustering, support vector machines, and ensemble methods.
Prerequisites: STAT 6420 or STAT 6220 or STAT 6315 or permission of department

STAT 6280
Credits: 3
Course Title: Applied Time Series Analysis
Course Description: Autoregressive, moving average, autoregressive-moving average, and integrated autoregressive-moving average processes, seasonal models, autocorrelation function, estimation, model checking, forecasting, spectrum, spectral estimators.
Prerequisites: [STAT 4360/6360 or STAT 4360E/6360E and (STAT 4230/6230 or STAT 6320 or STAT 6420)] or permission of department

STAT 6350
Credits: 3
Course Title: Applied Bayesian Statistics
Course Description: Introduction to theory and methods of the Bayesian approach to statistical inference and data analysis. Covers components of Bayesian analysis (prior, likelihood, posterior), computational algorithms, and philosophical differences among various schools of statistical thought.
Prerequisites: (STAT 4510/6510 and (STAT 4230/6230 or STAT 6420)) or permission of department

STAT 6420
Credits: 3
Course Title: Applied Linear Models
Course Description: Introduction to data analysis via linear models and logistic regression. Linear regression topics include estimation, inference, variable selection, diagnostics, and remediation. Basic design of experiments, analysis of variance, and logistic regression will also be covered, including an introduction to generalized linear models. Vector and matrix formulations are used throughout the course.
Prerequisites: Permission of department

STAT 6430
Credits: 3
Course Title: Design and Analysis of Experiments

Course Description: Theory and methods for constructing and analyzing designed experiments are considered. Basic concepts in design of experiments, analysis of covariance, completely randomized designs, randomized complete and incomplete block designs, row-column designs, repeated measures designs, factorial designs, split-plot experiments will be covered. Additional topics may include response surface modeling, mixture designs.
Prerequisites: STAT 6420 or permission of department

STAT 6510
Credits: 3
Course Title: Mathematical Statistics I
Course Description: Concepts and basic properties of some special probability distributions, independence, moment generating functions, sampling distributions of statistics, limiting distributions.
Prerequisites:

STAT 6620
Credits: 3
Course Title: Applied Categorical Data Analysis
Course Description: The methodology of categorical data analysis and its applications. The course covers descriptive and inferential methods for contingency tables, an introduction to generalized linear models, logistic regression, multinomial response models, regression for counts, and methods for categorical data from matched pairs.
Prerequisites or Corequisites: STAT 6420 and STAT 4510/6510

STAT 6800
Credits: 3
Course Title: Tools for Statistical Theory
Course Description: Provides preparation for graduate study in statistics by surveying topics in linear algebra and other areas chosen to strengthen students' analytical and mathematical skills.
Prerequisites: Permission of department

STAT 8000
Credits: 3
Course Title: Introductory Statistical Collaboration
Course Description:
Teaches students the communication skills necessary to successfully collaborate with non-statisticians in an interdisciplinary setting. Students will learn methods for conducting successful interactions with non-statisticians and will have opportunities to practice written and oral communication skills related to the application of statistics in other fields.
Prerequisites: STAT 6420

STAT 8060
Credits: 3

Course Title: Statistical Computing I
Course Description:
Tools and methods of statistical computing beginning with mathematical and computational underpinnings of statistical computation and progressing through Monte Carlo simulation, numerical linear algebra, optimization, numerical differentiation and integration, and simulation-based statistical algorithms. Students will learn methods, theory, and implementation via existing functions and their own code.
Prerequisites: (STAT 4510/6510 or STAT 6810) and STAT 6420

STAT 8070
Credits: 3
Course Title: Statistical Computing II
Course Description: Continuation of Statistical Computing I. Advanced statistical computing techniques will be covered. Topics may include advanced MCMC methods, Expectation-Maximization methods, machine-learning algorithms, constrained optimization, density estimation, nonparametric regression perfect sampling, data visualization, and parallel computing. Students will learn methods, theory, and implementation via existing functions and their own code.
Prerequisites: STAT 8060

STAT 8210
Credits: 3
Course Title: Multivariate: Theory and Methods
Course Description: An introduction to the theory and methodology of multivariate statistics for students with training in linear models and mathematical statistics. Topics include the multivariate normal distribution, one and two population inference on population mean vectors, MANOVA, principal component analysis, factor analysis, discrimination, classification, and canonical correlation.
Prerequisites: STAT 6420 and STAT 4520/6520

STAT 8230
Credits: 3
Course Title: Applied Nonlinear Regression
Course Description: Statistical modeling using nonlinear regression is considered. Topics include fixed-effects nonlinear regression models, nonlinear least squares, computational methods and practical matters, growth models, and compartmental models. Nonlinear mixed-effects models are discussed, including model interpretation, estimation and inference. Examples will be drawn from forestry, pharmaceutical sciences, and other fields.
Prerequisites: STAT 4230/6230 or STAT 6320 or STAT 6420 or permission of department

STAT 8260
Credits: 3
Course Title: Theory of Linear Models
Course Description:

Theory of the linear model is studied. Topics include a review of linear algebra; distribution theory; full and non-full rank linear models; ordinary and generalized least squares; maximum and restricted maximum likelihood estimation; prediction, inference, estimability, analysis of variance, restricted models, reparameterization, and mixed-effect models.
Prerequisites: STAT 6420 or permission of department
Prerequisites or Corequisites: STAT 6820 or STAT 4520/6520

STAT 8270
Credits: 3
Course Title: Spatial Statistics
Course Description:
Models and theories in spatial data, including geostatistics, lattice data, spatial point patterns, and space-time data. The course will focus on random field theory, various spatial regression models, model fitting, inferences and spatial prediction, with applications to agriculture, environmental sciences, forestry, and public health.
Prerequisites: STAT 4520/6520 and STAT 8260

STAT 8280
Credits: 3
Course Title: Time Series Analysis
Course Description:
Advanced topics in time series analysis and forecasting. Linear and nonlinear time series will be discussed. Topics include stationary processes, autocorrelation functions, various univariate time series models, forecasting, and multivariate time series. The focus is mostly on theoretical topics, but some applications are covered.
Prerequisites: STAT 6820 or STAT 8260

STAT 8290
Credits: 3
Course Title: Advances in Experimental Designs
Course Description: Covers state-of-the-art knowledge on selected topics such as factorial experiments, fractional factorials, incomplete block designs, orthogonal arrays, crossover designs, response surface methodology, mixture experiments, optimal design theory for linear and nonlinear models, and design construction techniques.
Prerequisites: STAT 8260 and STAT 6430

STAT 8330
Credits: 3
Course Title: Advanced Statistical Applications and Computing
Course Description: Advanced programming and implementation of modern statistical techniques using statistical software such as R. Topics include Monte Carlo simulations, resampling techniques, penalized regression, generalized linear models, robust methods, nonlinear regression, multiple testing adjustment, and smoothing techniques.
Prerequisites: [(STAT 6420 and (STAT 4360/6360 or STAT 4360E/6360E)] or permission of department

STAT 8620
Credits: 3
Course Title: Categorical Data Analysis and Generalized Linear Models
Course Description: Categorical data analysis and generalized linear models beginning with contingency tables and their analysis. Theory of generalized linear models will then be presented, followed by more detailed and application-oriented discussions of special cases, including logistic, log-linear models, and multinomial response models. Overdispersion is also discussed.
Prerequisites: STAT 8260 and (STAT 4520/6520 or STAT 6820)

STAT 8630
Credits: 3
Course Title: Mixed-Effect Models and Longitudinal Data Analysis
Course Description: Extensions of classical and generalized linear models with emphasis on longitudinal data analysis. Course will focus on linear mixed models, and marginal and mixed-effect versions of generalized linear models for longitudinal discrete data. Emphasis will be placed on the application of these models to analyze real data.
Prerequisites: STAT 8260
Prerequisites or Corequisites: STAT 8620

STAT 8920
Credits: 2
Course Title: Statistical Research and Professional Practice I
Course Description: Provides training in some of the skills, tools, and resources essential for conducting statistical research and for professional practice. Students will learn materials and methods for conducting statistical research as well as written and oral skills for communicating research. The course will also cover professional ethics for the statistician.
Prerequisites: Permission of department.

UNIVERSITY OF GEORGIA

Terry College of Business
Office of the Dean
417 Correll Hall
Athens, Georgia 30602
TEL 706-542-8100 | FAX 706-542-3835
busdean@uga.edu
terry.uga.edu

Benjamin C. Ayers, Dean
Earl Davis Chair in Taxation

MEMORANDUM

TO: S. Jack Hu, Senior Vice President for Academic Affairs and Provost

FROM: Benjamin C. Ayers, Terry College of Business Dean BCA

DATE: November 17, 2020

RE: Interdisciplinary Master's degree in Data Science

I am pleased to provide this letter of support on behalf of the Terry College of Business for the proposed M. S. in Data Science from Franklin College of Arts and Science's Departments of Statistics & Computer Science. This degree will be jointly administered by the Department of Computer Science and the Department of Statistics in the Franklin College of Arts and Sciences. The program of study in the proposed M. S. in Data Science places equal emphasis on both computer science and statistics in terms of the core courses while allowing students to select from a large array of elective courses in both disciplines to customize their focus. I have reviewed the proposal and concur that this degree does not present duplication of the Master of Science in Business Analytics (MSBA) degree offered by the Terry College of Business and agree that this would be a valuable program offering for students.

/abg

Documentation of Approval and Notification

Proposal: Major in Data Science

College: Franklin College of Arts and Sciences

Department: Department of Statistics and Department of Computer Science

Proposed Effective Term: Spring 2022

Department:

- Computer Science Department Head, Dr. Thiab R Taha, 12/1/21
- Statistics Department Head, Dr. T.N. Sriram, 12/1/20

School/College:

- Franklin College of Arts and Sciences Dean, Dr. Jean Martin-Williams, 2/15/21

Graduate School:

- Vice Provost for Graduate Education and Graduate School Dean, Dr. Ron Walcott, 4/14/21

Additional Support Letter:

- Terry College of Business Dean, Dr. Benjamin Ayers, 11/17/20