UNIVERSITY OF GEORGIA

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

University Council

March 16, 2018

UNIVERSITY CURRICULUM COMMITTEE – 2017-2018
Dr. Alison F. Alexander, Chair
Agricultural and Environmental Sciences – Dr. Elizabeth Little
Arts and Sciences -   Dr. Sujata Iyengar (Arts)
                      Dr. Mitch Rothstein (Sciences)
Business - Dr. Rich Gooner
Ecology – Dr. Sonia Altizer
Education – Dr. Morgan Faison
Engineering - Dr. Sudhagar Mani
Environment and Design – Professor Brad Davis
Family and Consumer Sciences - Dr. Patricia Hunt-Hurst
Forestry and Natural Resources - Dr. John C. Maerz
Journalism and Mass Communication – Dr. Jay Hamilton
Law – Professor Randy Beck
Pharmacy – Dr. Robin Southwood
Public and International Affairs - Dr. Robert Grafstein
Public Health – Dr. Anne Marie Zimeri
Social Work - Dr. David O. Okech
Veterinary Medicine - Dr. Kira L. Epstein
Graduate School – Dr. Amy Medlock
Ex-Officio - Provost Pamela S. Whitten
Undergraduate Student Representative – Mr. Max Harris
Graduate Student Representative – Ms. Johnita Daniel

Dear Colleagues:

The attached proposal from the Franklin College of Arts and Sciences to offer a new major in Data Science (B.S.) will be an agenda item for the March 23, 2018, Full University Curriculum Committee meeting.

Sincerely,

Alison Alexander

Alison F. Alexander, Chair
University Curriculum Committee

cc: Provost Pamela S. Whitten
        Dr. Rahul Shrivastav

Institution:	The University of Georgia

Date Completed at the Institution: February 21, 2018

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

Degree: Bachelor of Science

Major: Data Science

CIP Code: 30300101

Anticipated Implementation Date: Fall 2019

Delivery Mode (check the most appropriate delivery mode in the box below):

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 SACS-COC)	
Combination of off-campus and online (specify whether 50% or more is offered online for SACS-COC)	
Hybrid, combination delivery, but less than 50% of the total program is online based on SACS-COC	
Contractual Location (specify the location):	

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

**Departments:** Computer Science, Statistics

Departmental Contacts: Dr. Thiab Taha, Dr. T. N. Sriram

We request approval to establish an interdisciplinary undergraduate major in Data Science. It will reside in the Franklin College of Arts and Sciences and will be offered jointly by the Department of Computer Science and the Department of Statistics. Students will receive their degree in Data Science from the Franklin College.

1. **Rationale:** (Provide the rationale for proposing the new academic program.)

With the rapid advances in computing and storage capacity, we now live in the "Big Data" era, where data are collected in real time and analyzed in science, business, industry, and government. Data Science has emerged as a prominent field of study because it can provide valuable insights using Big Data for making informed decisions in business, human health, security, etc. 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.
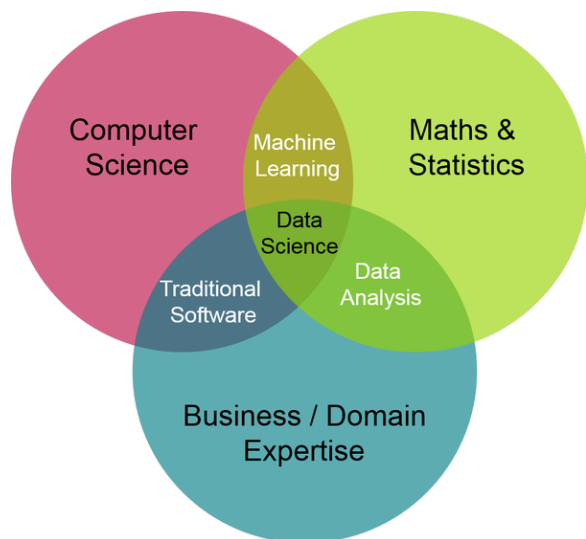
Data Science encompasses a wide range of concepts, methodologies, and algorithms involved in collecting, managing, visualizing, analyzing, and transforming *data* into information, knowledge-creation, and decision-making. The word *Data* refers to data collection, storage, and management and retrieval, while the word *Science* deals with modeling, analysis, inference, interpretation, and decision-making. Inherently, Data Science is a field of study that clearly rests at the intersection of Computer Science and Statistics.

Voluminous data is being regularly collected and analyzed in science, business, industry, as well as by government and society at large. This "data deluge" has 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, 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 "Big Data."

Currently, the Computer Science department offers courses related to Data Science and even a Certificate of Applied Data Science program. However, this certificate requires only minimal coursework in mathematics and statistics, which are fundamental topics for a rigorous Data Science program, for designing computational models in different fields, and for studying advanced topics in data analytics. The Terry College of Business has recently proposed a one-year (two-semester) Master of Science in Business Analytics (MSBA) program, which has been approved by the University and the Board of Regents. However, the MSBA program is aimed only at graduate students who are primarily interested in business applications. Thus, at the University of Georgia, there is no comprehensive undergraduate program that adequately trains students in all the areas involved in Data Science.

We explored the idea of offering Data Science as a track in the Statistics major or choosing a single home department for the Data Science major—either Computer Science or Statistics. We did not pursue either model 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 program will bridge key elements of Computer Science and Statistics into an interdisciplinary degree to develop future generations of data scientists. The program will also rely on prerequisites and electives taught in other units on campus, notably Mathematics and Management Information Systems, to provide students with a broad education and training.

Faculty members in Computer Science and Statistics have spent more than a year envisioning, planning, and designing the new program. It has the support of the dean of Franklin College, heads of the two departments, and members of the faculty in both departments. The proposal also includes letters of support from the director of Georgia Informatics Institutes for Research and Education, the dean of the Terry College of Business, and the head of the Department of Mathematics.

2. **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 of the scope of the institutional mission and sector, provide the compelling rationale for submission.

The development of this major 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 program 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. Undergraduate students in the Data Science program will benefit greatly from the synergistic activities organized by GII. This major also fits in with the trends in the nation. In order to meet the immediate demand for data scientists, many universities (University of Wisconsin, University of California Berkeley, Northwestern University, Carnegie Mellon University, Columbia University, and many more universities) across the U.S. have recently introduced Data Science programs at the graduate level. Furthermore, the July 2015 issue of AMSTAT NEWS (see page 22) published by the *American Statistical Association* carried an interesting article profiling five new undergraduate Data Science programs across the U.S. and Europe. More recently, three universities in the southeastern region of the U.S. (Auburn University, College of Charleston and Northern Kentucky University) have started bachelor's programs in Data Science. We expect the number of new undergraduate Data Science programs to increase substantially in the near future.

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

As industries and scientific research become more reliant on collecting and analyzing ever increasing amounts of data, the demand for data scientists is increasing dramatically. 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 undergraduate program in Data Science will provide necessary background in mathematics and build a strong foundation in Data Science, covering data structures, algorithms, and database management (courses in Computer Science) and data collection, data mining, machine learning, modeling, and inference (courses in Statistics). Some of the courses, such as data mining and machine learning, will be team-taught by faculty in Computer Science and Statistics to present perspectives from both disciplines. Students graduating with a degree in Data Science (B.S.) 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 experiential learning via a capstone course, which will focus on applying the acquired knowledge and skills in a real-world data analytics project. Upon graduation, students will be in high demand in the

One-Step Academic Program Proposal/Approval Form
RACAA Review July 16; Adopted August 30; Finalized October 3, 2016, USG System Office, MVMM

Page 3 of 52

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 will continue their education at the graduate level. Several avenues should be open to them, including MS/PhD in Data Science, Computer Science, Statistics, Management Information Systems or Industrial Engineering, as well as Master of Business Administration (MBA) and MSBA, which has been approved by the University and by the Board of Regents.

4. **Need:** Description of the justification of need for the program. (Explain in narrative form why the program is required to expand curricular 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.)

As indicated in John Tukey's book, 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 has 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 and will continue to fundamentally 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 and data on the internet and social media to sensor data captured by smart phones, automobiles, industrial systems, and environmental networks. As we have seen recently, governments are also using automated algorithms to mine large databases for the purposes of security. Examples abound in science: Sloan Digital Sky Survey, the Great Lakes Environmental Database, GenBank DNA sequences are driving scientific discovery and creation of new knowledge. Put simply, the enormous interest in Data Science now is a direct consequence of this confluence.

According to the McKinsey Report: there were 5 billion mobile phones in use in 2010; there are 30 billion pieces of content shared on Facebook every month; 15 out of 17 sectors in the U.S. have more data stored per company than the U.S. Library of Congress (which has 235 terabytes of data); and, even more stunning, there is a 40% projected growth in global data generated per year! At the same time, costs of data storage technologies are shrinking fast: we can buy a disk drive for $600 to store all of the world's music.

As a result of the technology changes, the McKinsey Report highlights that Big Data and Data Science will have a transformative potential in several domains in the coming years. As such, students graduating with a degree in Data Science will have a huge range of job opportunities. This is validated by the experience at North Carolina (NC) State University, which established a Master of Science in Analytics (MSA) in 2007. The NC State MSA job placement rates have exceeded 90percent by graduation over a period of nine consecutive years despite the cyclical job market. For example, the 113 students graduating in 2016 had about 280 reported job offers with an

average salary of about $93,000 (see MSA Employment Outcomes for more information). The 2016 Burtch Work Study reports that the median salary for Data Science level 1 professionals increased from $91,000 in 2015 to $97,000 in 2016, with Data Science managers earning well over $145,000 annually.

Other universities have established or are beginning to establish both graduate and undergraduate programs in Data Science. For example, University of California Irvine (UCI), Ohio State University (OSU), and University of Rochester (UR) have already established an undergraduate Data Science major. Based on our request, University of Rochester reported that they currently have over 50 declared majors (21 seniors, 23 juniors and 7 sophomores). In 2016, 2 students graduated with a Data Science degree from their program, whereas in 2017, they graduated 12 students. UCI and OSU provided more details about their Data Science program, given in the table below.

School	Semester-Year	No. Applications received	Admitted	Enrolled	Graduated
UCI	Fall 2015	20	15	12	
	Fall 2016	128	44	29	
	Fall 2017	281	141	111	
	Spring 2018				9-10 (expected)
	Spring 2019				20-25 (expected)
OSU	Fall 2014	40-50	10-20	14	
	Fall 2015	80-120	55-65	50-60	
	Fall 2016	80-120	55-65	50-60	
	Fall 2017	80-120	55-65	50-60	9
	Spring 2018				33 (expected)
	Spring 2019				40 (expected)

In each of these institutions, the Data Science program is thriving and becoming a major of choice. With the recent launching of GII, we believe the University of Georgia is especially well-positioned to offer a highly ranked undergraduate Data Science degree.

This new program will enhance existing collaborations between the Computer Science and Statistics departments that are spearheading the program. We also hope it will serve as a focal point for other units on campus to offer related courses as well as involve students in the program on internships and undergraduate research opportunities. There is considerable overlap in research interests in faculty and students in the two departments in the areas of computational biology, machine learning, biomedical informatics, bioinformatics, and data-intensive computational techniques.

If we wait, there are significant risks. First, 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. Second, there is a unique opportunity to make a big impact in the Data Science area for Georgia by establishing the GII as well as an undergraduate program. The GII institute can also serve undergraduates in Data Science by providing opportunities for undergraduate research, encouraging students to participate in research

seminars, etc.

5. **Demand:** Description of how the program demonstrates demand. (Explain in narrative form the data that supports demand for the program from existing and potential students and requests from regional industries.)

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. The 2011 McKinsey Report on Big Data projects that by 2018 there will be a shortage of talent necessary for organizations to take advantage of Big Data. The worldwide demand for individuals with data analysis skills will grow to almost 500,000, and by 2018 the United States alone could face a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million data-savvy managers and analysts to take full advantage of Big Data, analyze, and make decisions.

The White House has responded by launching a "*Big Data Research and Development Initiative*," to "*expand the workforce needed to develop and use Big Data technologies.*" This theme is reaffirmed by Thomas Davenport's article "*Data Scientist: The Sexiest Job in the 21st Century.*" There are already numerous job advertisements that specifically mention Data Science.
Partly fueled by these reports, Data Science programs have exploded in academic departments across the globe (e.g., U.S., Canada, Australia, New Zealand, and Europe) as universities find ways to meet the demand. There has been rapid growth of undergraduate programs at both research institutions and liberal arts colleges; see, for example, p. 22 of July 2015 issue of AMSTAT NEWS cited earlier. Therefore, there is a critical need to educate a new generation of UGA students who are rigorously trained in this multidisciplinary program, people who are comfortable with databases and algorithms as much as statistical modeling, inference, and visualization.

Recently, we conducted two short surveys, one during the summer 2017 semester and the other during the middle of fall 2017 semester, covering a wide spectrum of students. The first survey included incoming freshmen attending a Computational Science meeting held at the time of orientation on campus and students who took CSCI 1301 and STAT 4210 during summer 2017 semester. The second survey included students enrolled in the fall 2017 sections of CSCI 1302, CSCI 2610, CSCI 2720, STAT 4210, STAT 4220, STAT 4230, STAT 4510, and STAT 5010. The two surveys combined included a total of 642 respondents, consisting of students majoring in Computer Science (CS), Mathematics (MATH), Statistics (STAT), respectively; double majoring in CS and STAT, CS and MATH, MATH and STAT; and those majoring in other disciplines. The survey requested the respondents to rate their level of interest in pursuing a Data Science major on a scale of 0 to 5, with three key ratings defined as follows: "0= No Interest," "3=Indicated interest in a Data Science major," and "5=Change to the Data Science major."

Out of the 642 respondents, 92 of them said they would "consider changing" to the Data Science major. More importantly, 65 (out of 642) definitively said they would change to the Data Science major, which clearly establishes a very high level of interest in the new Data Science major. We conservatively estimate the projected enrollment to be approximately 10 new students per year.

We expect the new program to grow to about 150-200 students in four years (by 2022). This is significantly smaller than the current Computer Science program but closer to the number of declared majors in Statistics. In the long term, we expect the program to have a healthy growth, as other institutions also start establishing undergraduate and graduate Data Science programs and recruiters specifically start asking for Data Science majors. At that point in the future, the field will become an established major with a distinct identity.

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, female enrollments in statistics at the undergraduate level more closely mirror the overall campus gender distribution. We will send informative flyers about our Data Science program to public and private schools in Georgia, provide early exposure to Data Science to high school students through UGA Young Dawgs program, and hold multiple information sessions during new student orientations at UGA to make the program attractive to all high school students. We also plan on emphasizing the essential connections between Data Science and its applications to a variety of fields, such as biology, chemistry, medicine, social sciences, and marketing, which will also increase the demand for the new major.

6. **Duplication:** Description of how the program does not present duplication of existing academic offerings in the geographic area and within the system as a whole. 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.

On campus, the Computer Science department offers courses related to Data Science and a Certificate of Applied Data Science program, but these do not provide a comprehensive, rigorous training in Data Science that the new program hopes to accomplish. Also, since MSBA in the Terry College is a one-year program aimed only at graduate students who are primarily interested in business applications, our new undergraduate degree in Data Science does not present duplication of the MSBA program, as the two programs are demonstrably different in terms of the nature of students, course work, and more importantly, the emphasis in training. In Georgia, there are six universities currently offering master's in Data Science (and/or a closely related program). Kennesaw State University has even started a doctoral degree in Analytics and Data Science. The students entering these programs have an undergraduate or MS degree in one of the following fields: Computer Science, Mathematics, Statistics, Industrial Engineering, Information Systems, and Quantitative Finance. The intrinsic problem with this model is that it ignores the fact that an ideal data scientist is one who has a solid training in multiple fields: computer science, statistics, and mathematics. While these graduate programs address the problem by requiring their students to take undergraduate-level courses to fill the major gaps in educational training, they end up providing inadequate training in other key components or areas of Data Science. At best, these graduate programs provide a reasonable short-term solution to meet the immediate demand, but a more coherent and sustainable solution is to adequately prepare students at the undergraduate level.

Our new Data Science program is unique in Georgia in that it is designed to produce students with a deeper and broad intellectual understanding of both statistical and computing principles when

working with Big Data, with those principles being applicable to a variety of domains. This program will attract all UGA students who are interested using mathematics, statistics, and computing to solve a variety of Big Data problems. Further, the new program will help meet societal needs that otherwise would not be met, and more importantly, it increases the options for our undergraduate students. The McKinsey Report points to a shortfall of 140,000-190,000 data scientists by 2018, jobs that need to have deep analytical skills. This program should help prepare students to fill that need. Based on the current projections, there will be plenty of jobs in different sectors of the economy with different levels of experience required for them.

7. **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 Vice President for Academic Affairs) for the proposed academic program in the appendix.

8. **Forecast:** If this program was not listed on your academic forecast for the 2016 – 2017 academic year, 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.

9. **Admission Criteria:** List the admission criteria for the academic program.
   a) Include all required minima scores on standardized tests.

      No minimum scores are required.

   b) Include the required grade point average requirement.

      Students must complete one of the following courses with a grade of "C" (2.0) or better before entry into the Data Science major: CSCI 1301, CSCI 1302, or STAT 2010. To complete the Data Science major, students must complete each of the following courses with a grade of "C" (2.0) or better: CSCI 1301, CSCI 1302, CSCI (MATH) 2610, CSCI 2720, and STAT 2010.

10. **Curriculum (See the form below this series of questions and please complete.)**

a) 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.

   C.0. Summary of Course Requirements and Prerequisites:
   The course requirements for the B.S. in Data Science are listed below. The coursework consists of at least 67 semester hours with 18 hours of Foundation (Area VI) and at least 49 hours of major coursework.

- C.1. Foundation Courses (Area VI)

  CSCI 1302 (4) - Software Development
  CSCI 1360 (4) - Foundations for Informatics and Data Analytics
  CSCI 2150 (4) - Introduction to Computational Science
  CSCI 2720 (4) – Data Structures
  STAT 2010 (3) – Statistical Methods for Data Scientists (NEW)

- C.2. Major Required Courses (at least 37 credit hours)

  CSCI 3360 (4) - Data Science I
  CSCI 4260/6260 (4) - Data Security and Privacy (NEW)
  CSCI 4360/6360 (4) - Data Science II
  CSCI 4380/6380(4) or STAT 4250 (3) - Data Mining or Applied Multivariate Analysis
  and Statistical Learning (NEW)
  CSCI 4370/6370 (4) - Data Base Management
  MATH 3300 (3) – Applied Linear Algebra
  STAT 4220 (3) - Applied Experimental Designs
  STAT 4230/6230 (3) - Applied Regression Analysis
  STAT 4510/6510 (3) – Mathematical Statistics I
  STAT 4530 (3) – Statistical Inferences for Data Scientists (NEW)
  STAT (CSCI) 4990 (3) - Data Science Capstone (NEW)

- C.3 Major Elective Courses (choose 12 hours from the list below)

  BINF(PBIO) (3) – Bioinformatics Applications
  CSCI 3030 (3) - Computing, Ethics, and Society
  CSCI 4050/6050 (4) - Software Engineering
  CSCI 4150/6150(4) - Numerical Simulations in Science and Engineering
  CSCI 4210/6210 (4) - Simulation and Modeling
  CSCI 4470 (4)/6470 - Algorithms
  CSCI 4850/6850 (4) - Biomedical Image Analysis
  CSCI 5007/7007 (3) – Internship in Computer Science Business/Industry
  FINA 3001 (3) – Financial Management
  MARK 3001 (3) – Principles of Marketing
  MARK 4350 (3) - Marketing Analytics
  MARK 4650 (3) – Digital Marketing Analytics
  MATH (CSCI) 4690 (3) - Graph Theory
  MATH 4600 (3) – Probability
  MGMT 3001 (3) – Principles of Management
  MIST 5730 (3) - Advanced Data Management
  RMIN 4000 (3) - Risk Management and Insurance
  STAT 4240/6240 (3) - Sampling and Survey Methods
  STAT 4260/6260 (3) - Statistical Quality Assurance
  STAT 4280/6280 (3) - Applied Time Series Analysis

STAT 42906290 (3) - Nonparametric Methods
STAT 4360/6360 (3) – Statistical Software programming
STAT 4620/6620 (3) - Applied Categorical Data Analysis (NEW)
STAT 4710/6710 (3) - Introduction to Probability Theory I
STAT 4720/6720 (3) - Introduction to Probability Theory II
STAT 5700/7700 (3) – Internship in Statistics

b) Provide a sample program of study that includes the course prefixes, course numbers, and course titles and credit hour requirement for each course. Indicate the word "new" beside new courses.

Please see Appendix A for a sample program of study

c) List and reference all course prerequisites for required and elective courses within the program. Include the course prefixes, numbers, titles, and credit hour requirements.

BINF (PBIO) 4040/6040
Credits: 3
Course Title: Essential Biology for Quantitative Scientists
Prerequisites: CHEM 1211 or CHEM 1411 or CHEM 1311H or permission of department

CSCI 1302
Credits: 4
Course Title: Software Development
Prerequisite: CSCI 1301-1301L

CSCI 1360
Credits: 4
Course Title: Foundations for Informatics and Data Analytics
Prerequisites: CSCI1301 and CSCI 2610

CSCI 2150-2150L
Credits: 4
Course Title: Introduction to Computational Science
Prerequisites: MATH 1113 or permission of department

CSCI 2720
Credits: 4
Course Title: Data Structures
Prerequisites: CSCI 1730 and [CSCI (MATH) 2610 or CSCI 2611]

CSCI 3030
Credits: 3
Course Title: Computing, Ethics, and Society
Prerequisites: ENGL 1050H or ENGL 1102

CSCI 3360
Credits: 4
Course Title: Data Science I
Prerequisites: (CSCI 1301-1301L or CSCI 1360 or CSCI 1360E) and (MATH 2250 or MATH 2250E or CSCI 2150-2150L)

CSCI 4260 (NEW)
Credits: 4
Course Title: Data Security and Privacy
Prerequisites: CSCI 1302

CSCI 4360/6360
Credits: 4
Course Title: Data Science II
Prerequisites: CSCI 3360

CSCI 4370/6370
Credits: 4
Course Title: Data Base Management
Prerequisites: CSCI 2720

CSCI 4380/6380
Credits: 4
Course Title: Data Mining
Prerequisites: CSCI 2720

CSCI 4050/6050
Credits: 4
Course Title: Software Engineering
Prerequisites: CSCI 2720

CSCI 4150/6150
Credits: 4
Course Title: Numerical Simulations in Science and Engineering
Prerequisites: (MATH 2250 and CSCI 1301-1301L) or permission of department

CSCI 4210/6210
Credits: 4
Course Title: Simulation and Modeling
Prerequisites: CSCI 2720

CSCI 4470
Credits: 4
Course Title: Algorithms
Prerequisites: CSCI 2720

CSCI 4850/6850
Credits: 4
Course Title: Biomedical Image Analysis
Prerequisites: CSCI 2720 or permission of department

CSCI 5007/7007
Credits: 3 hours. Repeatable for maximum 12 hours credit.
Course Title: Internship in Computer Science Business/Industry
Prerequisites: Permission of department

FINA 3001
Credits: 3
Course Title: Financial Management
Prerequisites: ACCT 2101 or ACCT 2101E or ACCT 2101H or ACCT 1160 and MIST 2090 or MIST 2090E or MIST 2190H or CSCI 1100-1100L

MARK 3001
Credits: 3
Course Title: Principles of Marketing
Prerequisites: None

MARK 4350
Credits: 3
Course Title: Marketing Analytics
Prerequisites: MARK 3000 or MARK 3000E or MARK 3000H or MARK 3001 or MARK 3001H or MARK 3001E

MARK 4650
Credits: 3
Course Title: Integrated Marketing and Brand Communication
Prerequisites: MARK 3000 or MARK 3000E or MARK 3000H or MARK 3001 or MARK 3001H or MARK 3001E

MATH 3300
Credits: 3
Course Title: Applied Linear Algebra
Prerequisites: MATH 2260

MATH 4600/6600
Credits: 3
Course Title: Probability
Prerequisites: [(MATH 2270 or MATH 2500) and (MATH 2260 or MATH 3100)] or MATH 3510 or MATH 3510H

MATH (CSCI) 4690/6690

Credits: 3
Course Title: Graph Theory
Prerequisites: CSCI(MATH) 2610 or MATH 3200

MATH (BINF) 4780/ 6780
Credits: 3
Course Title: Mathematical Biology
Prerequisites: (MATH 2270 or MATH 2500 or MATH 3510 or MATH 3510H) and
[(MATH 4700/6700 or MATH 2700) and permission of department

MGMT 3001
Credits: 3
Course Title: Principles of Management
Prerequisites: (ACCT 2101 or ACCT 2101E or ACCT 2101H or ACCT 1160) and
(MIST 2090 or MIST 2090E or MIST 2190H or CSCI 1100-1100L)

MIST 5730
Credits: 3
Course Title: Advanced Data Management
Prerequisites: MIST 4600 and MIST 4610

PBIO (BINF) 4550/6550
Credits: 3
Course Title: Bioinformatics Applications
Prerequisites: BCMB 3600 or BCMB 3600H or GENE 3200-3200D

RMIN 4000
Credits: 3
Course Title: Risk Management and Insurance
Prerequisites: None

STAT 2010 (NEW)
Credit Hours: 3
Course Title: Statistical Methods for Data Scientists
Prerequisite: None

STAT 4220
Credit Hours: 3
Course Title: Applied Experimental Designs
Prerequisite: STAT 4110H or STAT 4210

STAT 4230/6230
Credit Hours: 3
Course Title: Applied Regression Analysis
Undergraduate Prerequisite: STAT 4210 or STAT 4110H

Graduate Prerequisite: STAT 6210 or STAT 6310 or STAT 6315 or permission of department

STAT 4240/6240
Credit Hours: 3
Course Title: Sampling and Survey Methods
Undergraduate Prerequisite: STAT 4210 or STAT 4110H
Graduate Prerequisite: STAT 6220 or STAT 6310 or STAT 6315 or permission of department

STAT 4250/6250 (NEW)
Credit Hours: 3
Course Title: Applied Multivariate Analysis and Statistical Learning
Undergraduate Prerequisite: STAT 4230 and (MATH 3000 or MATH 3300)
Graduate Prerequisite: (A course on linear algebra and either STAT 6420, STAT 6220, or STAT 6315) or permission of the department

STAT 4260/6260
Credit Hours: 3
Course Title: Statistical Quality Assurance
Undergraduate Prerequisite: STAT 4210 or STAT 4110H
Graduate Prerequisite: STAT 6220 or STAT 6310 or STAT 6315 or permission of department

STAT 4280/6280
Credit Hours: 3
Course Title: Applied Time Series Analysis
Undergraduate Prerequisite: STAT 4230/6230 and STAT 4360/6360
Graduate Prerequisite: [STAT 4360/6360 and (STAT 4230/6230 or STAT 6320 or STAT 6420)] or permission of department

STAT 4290/6290
Credit Hours: 3
Course Title: Nonparametric Methods
Undergraduate Prerequisite: STAT 4210 or STAT 4110H
Graduate Prerequisite: STAT 6220 or STAT 6310 or STAT 6315 or permission of department

STAT 4510/6510
Credit Hours: 3
Course Title: Mathematical Statistics I
Prerequisite: MATH 2270 or MATH 2500

STAT 4530 / CSCI 4530 (NEW)
Credit Hours: 3

Course Title: Estimation Methods for Data Science
Prerequisite: CSCI 3360, STAT4510

STAT 4620/6620 (NEW)
Credit Hours: 3
Course Title: Applied Categorical Data Analysis
Graduate Prerequisite: STAT 6420 and STAT 6510

STAT 4710/6710
Credit Hours: 3
Course Title: Introduction to Probability Theory I
Prerequisite: MATH 2500 or MATH 2270

STAT 4720/6720
Credit Hours: 3
Course Title: Introduction to Probability Theory II
Prerequisite: STAT 4710/6710

STAT 4990 / CSCI 4990 (NEW)
Credit Hours: 3
Course Title: Data Science Capstone Course
Prerequisite: CSCI 4360/6360, CSCI 4370/6370, STAT4220, STAT4230

STAT 5700/7700
Credit Hours: 3 hours. Repeatable for maximum 15 hours credit
Course Title: Internship in Statistics
Prerequisites: None

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 per the Academic and Student Affairs Handbook, Section 2.3.1.

Program of Study Form
(Modify appropriately for undergraduate versus graduate programs.)

The core curriculum for the University of Georgia can be found at
http://bulletin.uga.edu/GenEdCoreBulletin.aspx. Preferred courses for this major are included in the table.

Courses *(list acronym, number, and title)*	**Semester**	**Hours**
Area I: Foundation Courses		
ENGL 1101, English Composition I	First Year, Fall	3
ENGL 1102, English Composition II	First Year, Spring	3

MATH 2250, Calculus I for Science and Engineering	First Year, Fall	4
Area II: Physical and Life Sciences		
Life Science with a Laboratory	Second Year, Fall	4
Physical Science	Second Year, Spring	3
Area III: Quantitative Reasoning		
MATH 2260, Calculus II for Science Engineering	First Year, Spring	4
Area IV: World Languages and Culture		
World Language and Culture	Third Year, Fall	3
World Language and Culture	Third Year, Spring	3
World Language and Culture	Fourth Year, Spring	3
Area IV: Humanities and the Arts		
Humanities and the Arts	Fourth Year, Fall	3
Area V: Social Sciences		
HIST 2111, American History to 1865 or HIST 2112, American History since 1865	First Year, Fall	3
POLS 1101, American Government	Second Year, Fall	3
Social Science	Fourth Year, Fall	3
Area VI: CSCI 1302, Software Development	First Year, Spring	4
CSCI 1360, Foundations for Informatics and Data Analytics	First Year, Fall	4
CSCI 2150, Introduction to Computational Science	Second Year, Spring	4
CSCI 2720, Data Structures	Second Year, Spring	4
STAT 2010, Statistical Methods for Data Scientists	First Year, Spring	3
Major Area Courses – Common Curriculum CSCI 3360, Data Science	Third Year, Fall	4
CSCI 4260, Data Security and Privacy	Fourth Year, Fall	4
CSCI 4360, Data Science II	Third Year, Spring	4
CSCI 4370, Data Base Management	Third Year, Spring	4
CSCI 4380, Data Mining, or STAT 4250, Applied Multivariate Analysis and Statistical Learning	Third Year, Fall	4 or 3
MATH 3300, Applied Linear Algebra	Second Year, Spring	3
STAT 4220, Applied Experimental Designs	Second Year, Spring	3
STAT 4230, Applied Regression Analysis	Second Year, Spring	3
STAT 4510, Mathematical Statistics I	Third Year, Fall	3
STAT 4530, Estimation Methods for Data Science	Third Year, Spring	3
STAT(CSCI) 4990, Data Science Capstone	Fourth Year, Spring	2
Concentration		
Electives		

General Elective	Fourth Year, Spring	3
Major Elective	Fourth Year, Fall	3
Major Elective	Fourth Year, Fall	3
Major Elective	Fourth Year, Spring	3
Major Elective	Fourth Year, Spring	3
Multicultural Course	Third Year, Spring	3
Total Semester Credit Hours		
List below health and physical education, basic health, orientation, etc. per Board Policy 3.8.1		

e) 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, and telephone number. External reviewers must hold the rank of associate professor or higher in addition to other administrative titles.

   N/A

f) If internships, assistantships, or field experiences are required to complete the academic program, provide information documenting internship or field experience availability and how students will be assigned, supervised, and evaluated.

   Internships are optional but not required. Students may take one of the two following courses:

   CSCI 5007/7007: Internship in Computer Science Business/Industry
   STAT 5700/7700: Internship in Statistics

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

11. Waiver to Degree-Credit Hour (if applicable): State whether semester credit-hours exceed maximum limits for the academic program and provide a rationale.

   N/A

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

   All graduates earning the Data Science (B.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

(Goldman Sachs, AIG, Liberty Mutual) companies; pharmaceutical companies (Johnson & Johnson)]; government/national labs (NASA, NIST, DoD) or pursue graduate studies. All graduates will be able to:

- Develop software, 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, machine learning;
- Provide visualizations of the data and build statistical models to facilitate inference
- Interpret results of statistical analysis and assist decision makers

13. **Assessment and Quality**: Describe institutional assessments throughout the program to ensure academic quality, viability, and productivity as this relates to post-approval enrollment monitoring, degree productivity, and comprehensive program review.

All academic programs are reviewed annually to assess the program outcomes and student learning outcomes. Students completing Data Science (B.S.) are required to complete a capstone project, which requires applying a combination of concepts from computer science and statistics to do advanced data analytics in a domain area. The capstone 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.

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

   N/A

15. **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.

The University of Georgia has an existing base of students in the *Certificate of Applied Data Science* program offered by the Computer Science department. Our conservative enrollment projection assumes that in year 1, 2 of the existing certification students will shift into the new major and 8 new students will enter the major. We believe that the numbers shifted from other programs will slightly increase in subsequent years. Additionally, we conservatively estimate new enrollments at 9 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 56 majors by 2022.

If projections are not met, the directors of the program—heads of Computer Science and Statistics—along with the two respective undergraduate coordinators will develop a recruitment strategy to increase enrollment.

		First FY 2019	Second FY 2020	Third FY 2021	Fourth FY 2022
I. ENROLLMENT PROJECTIONS					
Student Majors					
Shifted from other programs		2	3	4	5
New to the institution		8	9	10	15
Total Majors **(Previous + new)**		10	22	36	56
Course Sections Satisfying Program Requirements					
Previously existing		0	4	8	15
New		1	0	3	2
Total Program Course Sections		1	4	11	17
Credit Hours Generated by Those Courses					
Existing enrollments		0	90	318-328	558-570
New enrollments		30	36	42	60
Total Credit Hours		30	126	360-370	618-630

16) Faculty

    a) Provide the total number of faculty members that will support this program: __**38**__

    b) Provide an inventory of faculty members directly involved with the administration and instruction of the program. Annotate in parentheses the person who holds the role of department chair. For each faculty member listed, provide the information below in tabular form. Indicate whether any positions listed are projected new hires and currently vacant.   (Multiple rows can be added to the table.) *Note: The table below is similar to the SACS-COC faculty roster form.*   *** (see narrative at the end of table)**

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 (Interim Head)	**Spring 2017** STAT 4110H, Honors Applied Statistics, 3.0 (UT) **Fall 2017** STAT 4/6280, Applied Time Series Analysis, 3.0 (UT)	Ph.D. Statistics, Michigan State University M.S. Statistics, University of Pune B.S. Statistics, Madras Christian College	3-6 credit hours/sem	Ph.D. thesis was in time series
Abhyuday Mandal	Associate Professor	**Spring 2017** STAT 4220, Applied Experimental Designs, 3.0 (UT) STAT 6430, Design and Analysis of Experiments, 3.0 (UT) **Summer 2017** STAT 4220, Applied Experimental Designs, 3.0 (UT) **Fall 2017** STAT 4220, Applied Experimental Designs, 3.0 (UT)	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/sem	Ph.D. thesis was in Design and Analysis of Experiments
Mark Werner	Senior Lecturer	**Spring 2017** MSIT 3000, Statistical Analysis for Business I, 3.0 (UT) [Taught 3 large sections of the course] **Summer 2017** MSIT 3000, Statistical Analysis for	Ph.D. Applied Mathematics, University of Colorado Denver M.S. Applied Mathematics, University of Stellenbosch B.S. Honors, Applied Mathematics,	18-24 credit hours per academic year	Ph.D. thesis was in Identification of Multivariate Outliers in Large Data Sets M.S. thesis was in Identifying Chaos in Experimental Data

		Business I, 3.0 (UT) STAT 4/6230, Applied Regression Analysis, 3.0 (UT) **Fall 2017** MSIT 3000, Statistical Analysis for Business I, 3.0 (UT) [Taught 3 large sections of the course]	University of Stellenbosch B.S. Applied Mathematics, University of Stellenbosch		
Ping Ma	Professor	**Spring 2017** STAT 4/6280, Applied Time Series Analysis, 3.0 (UT)	Ph.D. Statistics, Purdue University B.S. Mathematical Economics, Nankai University, China	3-6 credit hours/sem	Ph.D. thesis was in nonparametric estimation; Co-direct the Big Data Analytics Lab; Member of Georgia Informatics Institute; Adjunct Faculty Member of Computer Science
Wenxuan Zhong	Associate Professor	**Spring 2017** STAT 8460, Advanced Bioinformatics, 3.0 (GT)	Ph.D. Statistics, Purdue University B.S. Mathematical Economics, Nankai University, China	3-6 credit hours/sem	Ph.D. thesis was in dimension reduction; Co-direct the Big Data Analytics Lab; Founding Member of Georgia Informatics Institute
Pengsheng Ji	Assistant Professor	Spring 2017, STAT 6520, Mathematical Statistics, 3.0 (UT) Fall 2017, STAT 4/6230, Applied Regression Analysis, 3.0 (UT)	Ph.D. Statistics, Cornell University M.S. Statistics, Cornell University M.S. Statistics, Nankai University	3-6 credit hours/sem	Ph.D. thesis was in variable selection and nonparametric testing, which form the foundation of modern data science.

Name	Rank	Courses Taught	Education	Credit Hours	Research Interests
		STAT 6810, Probability Distributions, 3.0 (UT)	B.S. Mathematics, Nankai University		
Nicole Lazar	Professor	Fall 2017 STAT 5010/5020, Capstone Course, 3.0 (UT) Fall 2016 STAT 4/6230, Applied Regression Analysis, 3.0 (UT) STAT 6420, Applied Linear Models, 3.0 (UT)	Ph.D in Statistics, University of Chicago M.S. in Statistics, Stanford University B.A. in Statistics and Psychology, Tel Aviv University	6-9 credit hours per academic year	Ph.D in advanced likelihood methods; research in functional neuroimaging and other large data applications.
Paul Schliekelman	Associate Professor	Spring 2017 STAT 5010/5020, Capstone Course, 3.0 (UT) Fall 2017 STAT 6510, Mathematical Statistics I, 3.0 (UT) STAT 8060, Computing Techniques in Statistics I, 3.0 (UT)	Ph.D. in Biomathematics, North Carolina State University Master of Biomathematics, North Carolina State University B.S. in Physics, Iowa State University	9 credit hours per academic year	Ph.D. dissertation and research interests are interdisciplinary and involve large data sets.
Jaxk Reeves	Associate Professor	Spring 2017 STAT 4/6290,	Ph.D. in Statistics, University of	6 credit hours	Serving as Statistical

		Nonparametric Methods, 3.0 (UT) STAT 8000, Supervised Statistical Consulting, 3.0 (UT) Fall 2017 STAT 6420, Applied Linear Models, 3.0 (UT)	California, Berkeley B.S. in Applied Mathematics, Massachusetts Institute of Technology		Consulting Center (SCC) Director.
Megan Lutz	Lecturer	Spring 2017 MSIT 3000 (2 sections), Business Statistics, 3.0 (UT) STAT 4510, Mathematical Statistics, 3.0 (UT) STAT 6220, Introduction to Statistics II, 3.0 (UT) Fall 2017 STAT 4210 (2 sections), Statistical Methods, 3.0 (UT) STAT 4/6260 (Statistical Quality Assurance), 3.0 (UT) STAT 4510, Mathematical Statistics, 3.0 (UT)	Ph.D. in Psychology, Georgia Institute of Technology M.S. in Psychology Georgia Institute of Technology M.S. in Statistics Virginia Polytechnic Institute and State University M.S. in Applied Statistics Rochester Institute of Technology B.S. in Industrial Engineering Georgia Institute of Technology	24 credit hours per academic year	Ph.D. dissertation is in Psychology, which is an interdisciplinary area

			M.Sc. Mathematics (Num. Analysis), University of Jordan B.Sc. Mathematics, University of Jordan		
Hamid Arabnia	Professor (Graduate Coordinator)	**Spring 2017** CSCI 8990, Graduate Seminar, 1.0 (G) **Fall 2017** CSCI 4810/6810, Computer Graphics, 4.0 (UT/G)	Ph.D. Computer Science, University of Kent, Canterbury, England B.Sc. Honors Computer Science, Polytechnic of Wales, U.K.	5 credit hours/sem.	Ph.D. thesis: "Parallel Algorithms for Operations on Digitized Images"
Ismailcem Budak Arpinar	Associate Professor	**Spring 2017** CSCI 4350/6350, Global Information Systems, 4.0 (UT/G) **Fall 2017** CSCI, 8380, Advanced Topics in Information Systems, 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	Lecturer (Undergrad. Coordinator)	**Spring 2017** CSCI 1301, Introduction to Programing, 4.0 (UT) **Fall 2017**	Ph.D. Computer Science, University of Georgia B.S. Computer Science,	8 credit hours/sem.	Ph.D. dissertation: "A Regression Based System for Accurate Scalability Prediction on Large-Scale Machines"

Catherine Case	Lecturer	Spring 2017 STAT 2100H (2 sections), 4.0 (UT) STAT 4210 (3 sections), Statistical Methods, 3.0 (UT) Fall 2017 STAT 2100H (2 sections), 4.0 (UT) MSIT 3000 (2 sections), Business Statistics, 3.0 (UT)	Ph.D. in Mathematics & Statistics Education, University of Florida M.S. in Statistics, University of Florida B.S. in Mathematics and Spanish, University of Alabama at Birmingham	24 credit hours per academic year	Ph.D. dissertation is in statistics education, which is an interdisciplinary area.
Gauri S. Datta	Professor	Spring 2017 STAT 8530, Advanced Statistical Inference I, 3.0 (UT)	Ph.D. in Statistics, University of Florida M.Sc. in Statistics, University of Calcutta B.Sc. in Statistics, Ramakrishna Mission Residential College	9 credit hours per academic year	Research expertise in Bayesian Inference, and collaborates extensively with Census Bureau, CDC, and other Federal Agencies
Thiab Taha	Professor (Head)	**Spring 2017** CSCI 2150, Introduction to Computational Science, 4.0 (UT)	Ph.D. Applied Mathematics & Computer Science, Clarkson University	5 credit hours/year	Ph.D. dissertation: "On the Numerical and Analytical Aspects of Certain Nonlinear Evolution Equations"

		CSCI 1301, Introduction to Programming, 4.0 (UT)	College of Charleston		
Liming Cai	Professor	**Spring 2017** CSCI 4470/647, Algorithms, 4.0 (UT/G) **Fall 2017** CSCI 4470/647, Algorithms, 4.0 (UT/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/sem.	Ph.D. dissertation: "Limited Nondeterminism and Approximation" M.S. thesis: "A Directed Translation System for Programming Languages"
Michael Cotterell	Lecturer	**Spring 2017** CSCI 1730, Systems Programming, 4.0 (UT) **Fall 2017** CSCI 1302, Software Development, 4.0 (UT)	(IN PROGRESS) Ph.D. Computer Science, University of Georgia B.S. Computer Science, University of Georgia	12 credit hours/sem.	
Prashant Doshi	Professor	**Spring 2017** CSCI 8920, Decision Making, 4.0 (G) **Fall 2017** CSCI 4530/6530, Introduction to Robotics, 4.0 (UT/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	**Fall 2016** CSCI 2670, Introduction to Theoretical Computing, 4.0 (UT)	Ph.D. Computer Science, University of North Carolina, Chapel Hill	5 credit hours/sem.	Ph.D. dissertation: "EDF Scheduling on Heterogeneous Multiprocessors"

			M.S. Computer Science, University of North Carolina, Chapel Hill		M.S. thesis: "Using a Modified Size Measure to Guide the Search in the Ordered Semantic Hyper-Linking Theorem Prover"
			M.S. Mathematics, University of North Carolina, Chapel Hill		
			B.S. Mathematics, University of Maryland		
Bill Hollingsworth	Lecturer	**Spring 2017** CSCI 4690/6690, Graph Theory, 3.0 (UT/G) **Fall 2017** CSCI 2150, Introduction to Scientific 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"
Yi Hong	Assistant Professor	**Spring 2017** CSCI 4150/6150, Numerical Simulations in Science and Engineering, 4.0 (UT/G) **Fall 2017** CSCI 8955, Advanced Data Analytics: Statistical Learning and	Ph.D. Computer Science, University of North Carolina at Chapel Hill M.S. Computer Science, University of North Carolina at Chapel Hill M.S. Computer Science,	5 credit hours/sem.	Ph.D. dissertation: "Image and Shape Analysis for Spatiotemporal Data"

			Optimization, 4.0 (G)	Institute of Computing Technology, Chinese Academy of Sciences B.S. Computer Science, Wuhan University, China		
Maria Hybinette	Associate Professor	**Spring 2017** CSCI 4070/6070 Game Programming, 4.0 (UT/G) **Fall 2017** CSCI 4730/6730 Operating Systems, 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"
Manijeh Keshtgari	Lecturer	**Spring 2017** CSCI 4760, Computer Networks, 4.0 (UT **Fall 2017** CSCI 2610, Discrete Math, 4.0 (UT)	Ph.D. Computer Engineering, Sharif University of Technology M.S. Electrical & Computer Engineering, Colorado State University B.S. Computer Engineering, Shiraz University	12 credit hours/sem.	Ph.D. dissertation: 'Survivability of Networks M.S. thesis: "Performance Evaluation of FDDI-II Networks"
Krzysztof Kochut	Professor	**Spring 2017** CSCI 4570/6570, Compliers, 4.0 (UT/G) **Fall 2017**	Ph.D. Computer Science, Louisiana State University M.S. Computer Science,	5 credit hours/sem.	Ph.D. dissertation: "Resolution Proof Technique in Linear Temporal Logic" M.S. thesis: "Augmented

		CSCI 4050/6050, Software Engineering, 4.0 (UT/G)	University of Warsaw, Poland		Transition Network Grammar for Parsing of the Polish Language"
Kyu Hyung Lee	Assistant Professor	**Spring 2017** CSCI 4730/6730, Operating Systems, 4.0 (UT/G) **Fall 2017** CSCI 8240, Software Security and Cyber Forensics, 4.0 (UT)	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"
Jaewoo Lee	Assistant Professor	**Spring 2017** CSCI 3360, Data Science I, 4.0 (UT) **Fall 2017** 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"
Kang Li	Professor	**Fall 2017** CSCI 8250, Advanced Network Security, 4.0 (G)	Ph.D. Computer Science and Engineering, Oregon Health and Science University B.S. Computer Science and Engineering, Tsinghua University, China	5 credit hours/sem.	Ph.D. dissertation: "Modeling the Bandwidth Sharing Behavior of Congestion Controlled Flows"

Tianming Liu	Distinguished Research Professor	**Spring 2017** CSCI 4850/6850, Biomedical Image Analysis, 4.0 (UT/G) **Fall 2017** CSCI 8860, Biomedical Informatics, 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"
Sachin Meena	Lecturer	**Fall 2017** CSCI 2720, Data Structure, 4.0 (UT)	Ph.D. Computer Science, University of Missouri, B.S. Electrical Engineering, Indian Institute of Technology, New Delhi, India	12 credit hours/sem.	Ph.D. dissertation: "Interactive Image Segmentation Using a Few Points"
John Miller	Professor (Associate Head)	**Spring 2017** CSCI 8370, Advanced Database, 4.0 (G) **Fall 2017** CSCI 4370/6370, Database Management, 4.0 (UT/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"
Roberto Perdisci	Associate Professor	**Spring 2017**	Ph.D. Computer Engineering,	5 credit hours/sem.	Ph.D. thesis: "Statistical Pattern

		CSCI 4760/6760, Computer Networks, 4.0 (UT/G)	University of Cagliari, Italy M.S. Electronic Engineering, University of Cagliari, Italy		Recognition Techniques for Intrusion Detection in Computer Networks, Challenges and Solutions" M.S. thesis: "Alarm Clustering for Intrusion Detection Systems in Computer Networks"
Shannon Quinn	Assistant Professor	**Fall 2017** CSCI 4360/6360, Data Science II, 4.0 (UT/G) **Spring 2017** CBIO (CSCI) 4835/6835, Introduction to Computational Biology, 3.0 (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 2017** CSCI 4780/6780, Distributed Computing Systems, 4.0 (UT/G) **Fall 2017** CSCI 8780, Advanced Distributed Systems, 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 2017** CSCI 8950, Machine Learning, 4.0 (G) **Fall 2017** 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	**Spring 2017** CSCI 2720, Data Structures, 4.0 (UT) **Fall 2017** 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	**Fall 2017** CSCI 1730, Systems Programming, 4.0 (UT)	(IN PROGRESS) Ph.D. Computer Science, University of Georgia	12 credit hours/sem.	
Delaram Yazdansepas	Lecturer	**Fall 2017** CSCI 2610, Discrete Mathematics, 4.0 (UT)	(IN PROGRESS) Ph.D. Computer Science, University of Georgia B.S. Computer Engineering, Qazvin Azad University, Iran	12 credit hours/sem.	(IN PROGRESS) Ph.D. dissertation: "Human Activity Recognition using Wearable Sensors"

F, P: Full-time or Part-time: D, UN, UT, G: Developmental, Undergraduate Non-transferable, Undergraduate Transferable, Graduate

*The Department of Statistics and the Department of Computer Science are in the process of hiring 8 new faculty who will teach courses in this new Data Science program.

c) Explain how faculty workloads will be impacted by the proposed new program.

For Statistics: It is anticipated that the new co-director of the Data Science program (T. N. Sriram) will teach the new course—STAT 4530, Statistical Inferences for Data Scientists; Abhyuday Mandal will teach the new course—STAT (CSCI) 4990, Data Science Capstone and the existing course—STAT 4220, Applied Experimental Designs; and Catherine Case will teach the new course— STAT 2010, Statistical Methods for Data Scientists and Mark Werner will teach the existing course—STAT 4230/6230, Applied Regression Analysis.

For Computer Science: The Foundation Courses (C.1) are offered frequently. For the Major Related Courses (C.2), several faculty members are available to teach these courses: CSCI 1360 Foundations for Informatics and Data Analytics will be taught by Dr. Quinn; CSCI 2150 Introduction to Computational Science will be taught by Dr. Taha; CSCI 3360 Data Science will be taught by Drs. Hong, J. Lee, Miller and Quinn; CSCI 4260 Data Security and Privacy (NEW) will be taught by Dr. J. Lee; CSCI 4360 Data Science II will be taught by Drs. Hong, J. Lee, Miller and Quinn; CSCI 4380 Data Mining will be taught by Dr. Rasheed; CSCI 4370 Database Management will be taught by Drs. Arpinar and Miller.

In subsequent years, new faculty positions will be justified by increased enrollments.

d) Explain whether additional faculty will be needed to establish and implement the program. 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.

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 and have been authorized to hire at least one tenure-track faculty member and a non-tenure-track faculty member in spring 2018. In addition, from the 2018 Presidential Hiring Initiative for investing in the "Student Experience," Statistics has been authorized to hire a tenure-track faculty in statistical machine learning and a lecturer in the area of statistics and business statistics. Similarly, The Computer Science department has also added 3 top quality faculty in the last three years and received approval to hire 4 more

tenure-track assistant professors, starting August 2018. 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.

17) Fiscal and Estimated Budget
a) Describe the resources that will be used specifically for the program.

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

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

c) For Expenditures:

   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. We will also use the current infrastructure 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 offered that year. In our calculations, the average instructional cost for each new course in the major is taken to be $16,000, whereas for courses that are common to the Data Science (DS) major and the Statistics and Computer Science (CS) majors, we take the average instructional cost to be $4,000. 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 one new course (STAT 2010) for the major. In the second year, the expenditure is based on offering three courses (STAT 4230, MATH 3300, STAT 4220) that are common to Statistics, CS, and DS majors and the new course (STAT 2010) for those new students joining the major. In the third and fourth years, expenditures are determined based on similar calculations.

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).

Instructional time will be reallocated to teach new courses.

d) For Revenue:

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.

1. **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 major and the Statistics and Computer Sciences majors. Funds to teach the 6 new courses for the major will be reallocated using the new faculty lines which are already allocated to both departments.

2. **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.

3. **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 from the 2017-2018 rate of $4,776 for 6+ hours/semester. For each year, the anticipated number of students enrolled in the program during the fall and spring, respectively, is multiplied by $4,776 to obtain the total tuition per year. For example, we anticipate that 10 students will be enrolled in the program during the first year. Therefore, for Year 1, the total tuition amount is 10x$4,776 (fall) +10x$4,776 (spring) = $95,520. Similarly, we calculated the tuition amounts for years 2 to 4 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

e) 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. EXPENDITURES	First FY Dollars	Second FY Dollars	Third FY Dollars	Fourth FY Dollars
Personnel – reassigned or existing positions				
Faculty **(see 15.a.ii)**	$16,000	$28,000	$72,000	$100,000
Part-time Faculty **(see 15 a.ii)**				
Graduate Assistants **(see 15 a.ii)**				
Administrators**(see 15 a.ii)**				
Support Staff **(see 15 a.ii)**				
Fringe Benefits				
Other Personnel Costs				
Total Existing Personnel Costs	**$16,000**	**$28,000**	**$72,000**	**$100,000**

EXPENDITURES (Continued)				
Personnel – new positions (see 15 a.i)				
Faculty				
Part-time Faculty				
Graduate Assistants				
Administrators				
Support Staff				
Fringe Benefits				
Other personnel costs				
Total New Personnel Costs	$0	$0	$0	$0
Start-up Costs (one-time expenses) (see 15 a.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 15 a.i)				
Supplies/Expenses				
Travel				
Equipment				
Library/learning resources				
Other				
Total Recurring Costs				
GRAND TOTAL COSTS	$16,000	$28,000	$72,000	$100,000
III. REVENUE SOURCES				
Source of Funds				
Reallocation of existing funds (**see 15 b.i**)	$16,000	$28,000	$72,000	$100,000
New student workload				
New Tuition (**see 15 b.ii**)	$95,520	$210,144	$343,872	$534,912
Federal funds				
Other grants (**see 15 b.iv**)				
Student fees (**see 15 b.iii**) Exclude mandatory fees (i.e., activity, health, athletic, etc.).				
Other (**see 15 b.v**)				

New state allocation requested for budget hearing				
GRAND TOTAL REVENUES	**$111,520**	**$238,144**	**$415,872**	**$634,912**
Nature of Revenues				
Recurring/Permanent Funds				
One-time funds				
Projected Surplus/Deficit (Grand Total Revenue – Grand Total Costs) (see 15 c.i. & c.ii).	**$95,520**	**$210,144**	**$343,872**	**$534,912**

18) 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)**
	15	Classrooms	40	37,500
	2	Labs (dry)	50	6,250
		Labs (wet)		
		Meeting/Seminar Rooms		
		Offices		
		Other (specify)		
Total Assignable Square Feet (ASF)				**43,750**

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	

One-Step Academic Program Proposal/Approval Form
RACAA Review July 16; Adopted August 30; Finalized October 3, 2016, USG System Office, MVMM

Page 40 of 52

APPENDIX

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

	Corse Number	Course Title	Hours
First Year Fall	CSCI 1360	Foundations for Informatics and Data Analytics	4
	ENGL 1101	English Composition I	3
	FYOS 1000	First Year Odyssey Seminar	1
	MATH 2250	Calculus I for Science and Engineering	4
	HIST 2111 or	American History to 1865	3
	HIST 2112	American History since 1865	
		Total Credit Hours	15
First Year Spring	CSCI 1302	Software Development	4
	ENGL 1102	English Composition II	3
	MATH 2260	Calculus II for Science Engineering	4
	PEDB	Physical Education	1
	STAT 2010	Statistical Methods – NEW	3
		Total Credit Hours	15
Second Year Fall	Area II	Life Science with Laboratory	4
	General Elective		3
	POLS 1101	American Government	3
	STAT 4230	Applied Regression Analysis	3
		Total Credit Hours	13
Second Year Spring	Area II	Physical Science without Laboratory	3
	CSCI 2150-2150L	Introduction to Computational Science	4
	CSCI 2720	Data Structures	4
	MATH 3300	Applied Linear Algebra	3
	STAT 4220	Applied Experimental Designs	3
		Total Credit Hours	17
Third Year Fall	Area IV	World Language and Culture	3
	CSCI 3360	Data Science	4
	STAT 4250 or	Applied Multivariate Analysis and Statistical	3 or
	CSCI 4380	Learning or Data Mining	4
	STAT 4510	Mathematical Statistics I	3
		Total Credits	13

Third Year Spring	Area IV	World Language and Culture	3
	CSCI 4360	Data Science II	4
	CSCI 4370	Data Base Management	4
	Multicultural	Franklin College Multicultural Course	3
	STAT 4530	Estimation Methods for Data Science (NEW)	3
		Total Credits	17
Fourth Year Fall	Area V	Social Science	3
	CSCI 4260	Data Security and Privacy	4
	Area IV	Humanities and the Arts	3
	Major Elective		3
	Major Elective		3
		Total Credit Hours	16
Fourth Year Spring	Area IV	World Language and Culture	3
	General Elective		3
	Major Elective		3
	Major Elective		3
	STAT/CSCI 4990	Data Science Capstone	3
		Total Credit Hours	15

Appendix B: Course Titles, Credits, Descriptions

BINF (PBIO) 4040/6040
Credits: 3
Course Title: Essential Biology for Quantitative Scientists
Course Description: The essential elements of biology necessary for a scientist with a background in the quantitative sciences to begin working in the biological sciences. Core biological concepts will be presented with the goal of getting computer scientists, physicists, and mathematicians started in understanding biology and finding productive areas of research inquiry.
Prerequisites: CHEM 1211 or CHEM 1411 or CHEM 1311H or permission of department

CSCI 1302
Credits: 4
Course Title: Software Development
Course Description: Software development techniques in an object-oriented computer language. An intermediate programming course emphasizing systems methods, top-down

design, testing, modularity, and structured techniques. Applications from areas of numeric and non-numeric processing and data structures.
Prerequisite: CSCI 1301-1301L

CSCI 1360
Credits: 4
Course Title: Foundations for Informatics and Data Analytics
Course Description: An introduction to concepts in scientific programming and Data Science using the Python language. Students are given hands-on opportunities to learn techniques applicable to quantitative analyses across a broad range of fields. Core programming concepts are taught in tandem with real-world applications.
Prerequisites: MATH 1113 or MATH 1113E

CSCI 2150-2150L
Credits: 4
Course Title: Introduction to Computational Science
Course Description: Basic topics of scientific computing that are necessary for science and engineering students. Solving mathematical problems by different numerical methods. Quantitative reasoning concepts will be emphasized in comparing and verification of the correctness of the solutions. Mathematical software packages will be used. This course is intended for freshman and sophomore students.
Prerequisites: MATH 1113 or permission of department

CSCI 2720
Credits: 4
Course Title: Data Structures
Description: The design, analysis, implementation, and evaluation of the fundamental structures for representing and manipulating data: lists, arrays, trees, tables, heaps, graphs, and their memory management.
Prerequisites: CSCI 1730 and [CSCI (MATH) 2610 or CSCI 2611]

CSCI 3030
Credits: 3
Course Title: Computing, Ethics, and Society
Course Description: Introduction to social and ethical issues relating to computer science and information technology. Topics include intellectual property, open source software, the digital divide, globalization, and professional ethics. Students should have a working knowledge of personal computing.
Prerequisites: ENGL 1050H or ENGL 1102

CSCI 3360
Credits: 4
Course Title: Data Science I
Course Description: A rigorous overview of methods for text mining, image processing, and scientific computing. Core concepts in supervised and unsupervised analytics, dimensionality reduction, and data visualization will be explored in depth.

Prerequisites: (CSCI 1301-1301L or CSCI 1360 or CSCI 1360E) and (MATH 2250 or MATH 2250E or CSCI 2150-2150L)

CSCI 4260 (NEW)
Credits: 4
Course Title: Data Security and Privacy
Course Description: This course will examine security and privacy issues related to protecting personal data in various environments (for examples: cloud computing, smart grid, and internet of things), cover the fundamentals and principles of data security and privacy, and explore computational and statistical techniques for constructing secure and private systems.
Prerequisites: CSCI 1302

CSCI 4360/6360
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
Prerequisites: CSCI 3360

CSCI 4370/6370
Credits: 4
Course Title: Data Base 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.
Prerequisites: CSCI 2720

CSCI 4380/6380
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.
Prerequisites: CSCI 2720

CSCI 4050/6050
Credits: 4
Course Title: Software Engineering
Course Description: Full cycle of a software system development effort, including requirements definition, system analysis, design, implementation, and testing. Special emphasis is placed on system analysis and design. The design phase includes

development of a user interface. A large term project incorporates the full software life cycle.
Prerequisites: CSCI 2720

CSCI 4150/6150
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 4210/6210
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.
Prerequisites: CSCI 2720

CSCI 4470
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.
Prerequisites: CSCI 2720

CSCI 4850/6850
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.
Prerequisites: CSCI 2720 or permission of department

CSCI 5007/7007
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.

Prerequisites: Permission of department

FINA 3001
Credits: 3
Course Title: Financial Management
Course Description: The basic concepts and analytical tools of finance in both corporate finance and investments. Topics include risk and return, financial institutions, efficient markets, valuation theory, capital budgeting, portfolio theory, cost of capital, and international finance. For non-Terry College of Business students.
Prerequisites: (ACCT 2101 or ACCT 2101E or ACCT 2101H or ACCT 1160) and (MIST 2090 or MIST 2090E or MIST 2190H or CSCI 1100-1100L)

MARK 3001
Credits: 3
Course Title: Principles of Marketing
Course Description: Examination of the ways organizations satisfy consumer and organizational needs and wants for products and services. Emphasis is placed on how the elements of the marketing mix are used to provide customer satisfaction in the marketing of goods and services, both domestically and internationally. For non- Terry College of Business students.

MARK 4350
Credits: 3
Course Title: Marketing Analytics
Course Description: A practitioner-oriented introduction to the methods of marketing analytics. Class examples and exercises are drawn from real-world applications, and students will be actively engaged in executing the methodologies using relevant software.
Prerequisites: MARK 3000 or MARK 3000E or MARK 3000H or MARK 3001 or MARK 3001H or MARK 3001E

MARK 4650
Credits: 3
Course Title: Integrated Marketing and Brand Communication
Course Description: Examination of how businesses and brands measure the effectiveness of digital marketing strategies. Emphasis is placed on using data to both evaluate a company's level of digital marketing sophistication and to recommend digital marketing strategies.
Prerequisites: MARK 3000 or MARK 3000E or MARK 3000H or MARK 3001 or MARK 3001H or MARK 3001E

MATH 3300
Credits: 3
Course Title: Applied Linear Algebra
Course Description: Linear algebra from an applied and computational viewpoint. Linear equations, vector spaces, linear transformations; linear independence, basis, dimension;

orthogonality, projections, and least squares solutions; eigenvalues, eigenvectors, singular value decomposition. Applications to science and engineering.
Prerequisites: MATH 2260

MATH 4600/6600
Credits: 3
Course Title: Probability
Course Description: Discrete and continuous random variables, expectation, independence and conditional probability; binomial, Bernoulli, normal, and Poisson distributions; law of large numbers and central limit theorem.
Prerequisites: [(MATH 2270 or MATH 2500) and (MATH 2260 or MATH 3100)] or MATH 3510 or MATH 3510H

MATH (CSCI) 4690/6690
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.
Prerequisites: CSCI (MATH) 2610 or MATH 3200

MATH (BINF) 4780/ 6780
Credits: 3
Course Title: Mathematical Biology
Course Description: Mathematical models in the biological sciences, systems, phase-plane analysis, diffusion, convective transport, bifurcation analysis. Possible applications will include population models, infectious disease and epidemic models, acquired immunity and drug distribution, tumor growth, and analysis of arterial flow dynamics.
Prerequisites: (MATH 2270 or MATH 2500 or MATH 3510 or MATH 3510H) and [(MATH 4700/6700 or MATH 2700) and permission of department

MGMT 3001
Credits: 3
Course Title: Principles of Management
Course Description: Management functions and processes as applied to organizations and to individuals in organizations. Topics to be covered include strategy, job and organization design, ethics and social responsibility, diversity, global influences, leadership, motivation, human resource management, and organizational change. Other emerging managerial issues may be introduced.
Prerequisites: (ACCT 2101 or ACCT 2101E or ACCT 2101H or ACCT 1160) and (MIST 2090 or MIST 2090E or MIST 2190H or CSCI 1100-1100L)

MIST 5730
Credits: 3
Course Title: Advanced Data Management

Course Description: SGML, XML and associated technologies, XML implementations, XML tools, XML data management, business uses and applications of XML, data synchronization technologies, and managing data synchronization system.
Prerequisites: MIST 4600 and MIST 4610

PBIO (BINF) 4550/6550
Credits: 3
Course Title: Bioinformatics Applications
Course Description: State-of-the-art computational analyses of genome, DNA, RNA, and protein sequences will be presented, including programs for analyzing these data and the underlying analysis methods. Topics include sequence and structure databases; sequence assembly; sequence alignment; evolutionary analyses; gene function prediction; genome annotation; and applications for medical, agricultural and environmental genomics.
Prerequisites: BCMB 3600 or BCMB 3600H or GENE 3200-3200D

RMIN 4000
Credits: 3
Course Title: Risk Management and Insurance
Course Description: Risk identification, risk evaluation, and risk treatment methods. Consideration is given to both business and personal risks, with an emphasis on insurance as a risk management tool. Designed for non-majors as well as a basis for more advanced courses.
Prerequisites:

Course ID: STAT 2010 (NEW)
Credit Hours: 3
Course Title: Statistical Methods for Data Scientists
Course Description: In-depth introductory statistical methods, focusing on inference, alignment between study design and conclusions, and real-world decision making. Includes parametric and non-parametric approaches to one-and two-sample inference for means and proportions, Type I and II errors, power; chi-squared tests, and simple regression. Course will be implemented in R.
Prerequisite:

Course ID: STAT 4220
Credit Hours: 3
Course Title: Applied Experimental Designs
Course Description: Constructing and analyzing statistical experimental designs; blocking, randomization, replication and interaction; complete and incomplete block designs; factorial experiments; repeated measures; confounding effects; orthogonal arrays; computer experiments and simulations; design and analysis for generalized linear models. (NEW Course Description)
Prerequisite: STAT 4110H or STAT 4210

Course ID: STAT 4230/6230
Credit Hours: 3

Course Title: Applied Regression Analysis
Course Description: Applied methods in regression analysis with implementation in R. Topics include linear regression with mathematical examination of model assumptions and inferential procedures; multiple regression and model building including collinearity, variable selection and inferential procedures; ANOVA as regression analysis; analysis of covariance; diagnostic checking techniques; generalized linear models including logistic regression. (NEW Course Description)
Undergraduate Prerequisite: STAT 4210 or STAT 4110H
Graduate Prerequisite: STAT 6210 or STAT 6310 or STAT 6315 or permission of department

Course ID: STAT 4240/6240
Credit Hours: 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.
Undergraduate Prerequisite: STAT 4210 or STAT 4110H
Graduate Prerequisite: STAT 6220 or STAT 6310 or STAT 6315 or permission of department

Course ID: STAT 4250/6250 (NEW)
Credit Hours: 3
Course Title: Applied Multivariate Analysis and Statistical Learning
Course description: An introduction to 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.
Undergraduate Prerequisite: STAT 4230 and (MATH 3000 or MATH 3300)
Graduate Prerequisite: (A course on linear algebra and either STAT 6420, STAT 6220, or STAT 6315) or permission of the department

Course ID: STAT 4260/6260
Credit Hours: 3
Course Title: Statistical Quality Assurance
Course description: Basic graphical techniques and control charts. Experimentation in quality assurance. Sampling issues. Other topics include process capability studies, error analysis, SPRT, estimation and reliability.
Undergraduate Prerequisite: STAT 4210 or STAT 4110H
Graduate Prerequisite: STAT 6220 or STAT 6310 or STAT 6315 or permission of department

Course ID: STAT 4280/6280
Credit Hours: 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.
Undergraduate Prerequisite: STAT 4230/6230 and STAT 4360/6360
Graduate Prerequisite: [STAT 4360/6360 and (STAT 4230/6230 or STAT 6320 or STAT 6420)] or permission of department

Course ID: STAT 4290/6290
Credit Hours: 3
Course Title: Nonparametric Methods
Course description: Techniques and applications of nonparametric statistical methods, estimates, confidence intervals, one sample tests, two sample tests, several sample tests, tests of fit, nonparametric analysis of variance, correlation tests, chi-square test of independence and homogeneity, sample size determination for some nonparametric tests.
Undergraduate Prerequisite: STAT 4210 or STAT 4110H
Graduate Prerequisite: STAT 6220 or STAT 6310 or STAT 6315 or permission of department

Course ID: STAT 4510/6510
Credit Hours: 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.
Prerequisite: MATH 2270 or MATH 2500

Course ID: STAT 4530 / CSCI 4530 (NEW)
Credit Hours: 3
Course Title: Estimation Methods for Data Science
Course description: Mathematical and computational approaches to estimation and inference from frequentist and Bayesian perspectives. Sampling distributions; maximum likelihood estimation; computational maximization of likelihoods, including grid search, Newton-Raphson methods; likelihood ratio tests. Simulations of power and error rates. Introduction to Bayesian inference; prior and posterior distributions; model building; sampling from the posterior distribution; MCMC algorithms.
Prerequisite: CSCI 3360, STAT4510

Course ID: STAT 4620/6620 (NEW)
Credit Hours: 3
Course Title: Applied Categorical Data Analysis
Course description: This is an introduction to 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.

Undergraduate Prerequisite: STAT 4220 and STAT 4230 and STAT 4510
Graduate Prerequisite: STAT 6420 and STAT 6510

Course ID: STAT 4710/6710
Credit Hours: 3
Course Title: Introduction to Probability Theory I
Course description: Probability axioms, combinatorial analysis, random variables, univariate and multivariate distributions, expectations, conditional distributions, independence, and laws of large numbers.
Prerequisite: MATH 2500 or MATH 2270

Course ID: STAT 4720/6720
Credit Hours: 3
Course Title: Introduction to Probability Theory II
Course description: Central limit theorems, random walks, Markov chains and processes, Brownian motion, branching and renewal processes, diffusion processes and queueing processes and applications.
Prerequisite: STAT 4710/6710

Course ID: STAT 4990 / CSCI 4990 (NEW)
Credit Hours: 3
Course Title: Data Science Capstone Course
Course description: Provides students with an exposure to advanced methods and technologies in Data Science including data acquisition, data quality, Big Data management and analytics, data mining, data security and privacy, and introduces the students to Data Science experience to a real-world problem. In addition, effective oral and written communication of technologies, methods and results are emphasized.
Prerequisite: CSCI 4360/6360, CSCI 4370/6370, STAT4220, STAT4230

STAT 5700/7700
Credit Hours: 3 hours. Repeatable for maximum 15 hours credit
Course Title: Internship in Statistics
Course description: Supervised practicum in a government agency or industry site
Prerequisites: None

Approvals on File

Proposal: New Major in Data Science (B.S.)

College: Franklin College of Arts and Sciences

Department(s): Department of Computer Science
Department of Statistics

Proposed Effective Term: Fall 2018

Department:

- Computer Science Department Head, Dr. Thiab Taha, 12/4/2017
- Statistics Interim Department Head, Dr. T.N. Sriram, 12/4/2017

School/College:

- Franklin College of Arts and Sciences Associate Dean, Dr. Jean Martin-Williams, 2/7/2018