

Routing for On-Campus Approval of Degree Program Actions

Type of Action: Enter "X" for Action Type(s) and list Title and Prefix(s) as indicated

☒ New Degree Program	Proposed Program Title <u>Foundations of Data Science</u>
☐ New Certificate Program	Proposed Certificate Program Title _____
☐ New Minor Program	Proposed Minor Program Title _____
☐ Change in Degree Program Title	Current Degree Program Title _____
☐ Change in Certificate Program Title	Current Certificate Program Title _____
	Proposed Certificate Program Title _____
☐ Change in Minor Program Title	Current Minor Program Title _____
	Proposed Minor Program Title _____
☐ Change in Course Prefix	Current Course Prefix _____ Proposed Course Prefix _____
☐ Program Discontinuation	

Proposed Effective Date Fall 2019 Program Contact: Pierre Gremaud (Math); George Rouskas (Computer Science); Wenbin Lu (Statistics)

Proposed CIP Code (see <https://nces.ed.gov/ipeds/cipcode/default.aspx?y=55>): 30.3001

Routing of Action: Indicate date when the following occurs

Completed Request to Plan and 1-page Concept Paper

9/22/10 Council of Dean's- Approval to Plan

Completed Proposal

3/30/18; 4/2/18 Department Head endorses*

3/29/18; 4/2/18 College Curriculum Committee (undergraduate or graduate) recommends*

4/2/18 College Dean endorses*

Proposal moves to Undergraduate or Graduate office for routing

☐ Recommended by Vice Provost, DELTA, if applies*

☐ Substantive Change Review Team (SCRT) informed

☐ _____ SACS liaison prepares letter for Chancellor to send to SACS, if notification is required

☐ University Courses & Curriculum Committee **or** Administrative Board of the Graduate School recommends

☐ Associate Deans Council **or** Graduate Operations Council informed

☐ Dean (Graduate School or DASA) approves*

Proposal move to the Executive Vice Chancellor Provost's office for routing

☐ Vice-Provosts informed

☐ Deans' Council recommends*

☐ Executive Vice Chancellor and Provost approves*

☐ Chancellor's Executive Officer's (EOM) recommend

☐ University Council informed

☐ Board of Trustees subcommittees recommend

☐ Chancellor approves*

☐ Accreditation Liaison notifies SACS, if applicable

☐ Submitted to UNC System Office by Provost's Office

* Signature is required on the signature page for the action

Master of Science in Foundations of Data Science
North Carolina State University

This request has been reviewed and approved by the appropriate campus committees and authorities.

Endorsed By:

Laurie Williams

Laurie Williams _____ 3/30/18
Head, Department of Computer Science (Printed Name and Signature) Date

Alina Chertock

Head, Department of Mathematics (Printed Name and Signature) 4/2/2018
Date

Leonard A. Stefanski

Head, Department of Statistics (Printed Name and Signature) 4-2-18
Date

Recommended By:

Douglas S. Reeves

Douglas S. Reeves 03/29/2018

Chair, College Graduate Studies Committee COE (Printed Name and Signature)
Date

ALUN L LLOYD

Chair, College Graduate Studies Committee COS (Printed Name and Signature) 4/2/18
Date

Endorsed By:

Louis A. Martin-Vega

College Dean COE (Printed Name and Signature) 4/3/18
Date

MC McGahan

College Dean COS (Printed Name and Signature) 4/2/18
Date

Recommended By:

Thomas K. M:110111

Vice Provost, DELTA (if DE degree) (Printed Name and Signature) 4/12/18
Date

Approved By:

Dean of the Graduate School (Printed Name and Signature) Date

Recommended By:

Dean's Council (Printed Name and Signature) Date

Approved By:

Executive Vice Chancellor and Provost

(Printed Name and Signature)

Date

Approved By:

Chancellor

(Printed Name and Signature)

Date



Letter of Intent to Develop New Academic Degree Program

The following approvals must be obtained prior to sending the Letter of Intent to Develop a New Academic Degree Program to the UNC System Office.

Institution North Carolina State University _____

Degree Program Title (e.g. M.A. in Biology) M.S. in Foundations of Data Science _____

Reviewed and Approved By (Name and title only. No signature required in this section.)

Check box to indicate participation in review. (Provost is required.)

- ☐ **Provost:**
- ☐ **Faculty Senate Chair (as appropriate):**
- ☐ **Graduate Council (as appropriate):**
- ☐ **Undergraduate or Graduate Dean (as appropriate):**
- ☐ **Academic College Dean:**
- ☐ **Department Chair:**
- ☐ **Program Director/Coordinator:**

New Academic Proposal Process

New academic programs are initiated and developed by the faculty members. Approval of the Letter of Intent to Develop a New Academic Degree Program must be obtained from department chairs and college deans or equivalent administrators before submission to the UNC System Office review.

Directions: Please provide a succinct, yet thorough response to each section. Obtain the Provost's signature and submit the proposal via the PREP system to the UNC System Vice President for Academic Programs, Faculty, and Research, for review and approval by the UNC System Office. Once the Letter of Intent to Develop is approved, the institution can begin work on the formal Request to Establish a New Degree Program.

Letter of Intent to Develop a New Academic Degree Program

Institution	North Carolina State University
Joint Degree Program (Yes or No)? If so, list partner campus.	No
Degree Program Title (e.g. M.A. in Biology)	M.S. in Foundations of Data Science
CIP Code and CIP Title (May be found at National Center for Education Statistics)	<u>30.3001</u>
Require UNC Teacher Licensure Specialty Area Code (Yes or No). If yes, list suggested UNC Specialty Area Code(s).	No
Proposed Delivery Mode (campus, online, or site-based distance education). Add maximum % online, if applicable.	The degree will be offered both on-campus and online. Online delivery will be asynchronous with on-campus courses.
Proposed Term to Enroll First Students (e.g. Spring 2019)	Fall 2019
List other programs in the UNC System (may be found at UNC System website)	NC State: Master of Science in Analytics (MSA); UNC Charlotte: Professional Science Master's in Data Science and Business Analytics (DBSA).

SACSCOC Liaison Statement: *(Provide a brief statement from the University SACSCOC liaison regarding whether the new program is or is not a substantive change.)*

Program Summary: *(Briefly describe the proposed program and summarize the overall rationale.)*
Maximum of 1,000 words.

Include the following in your narrative:

- Ways in which the proposed program is distinct from others already offered in the UNC System (use the 4-digit CIP as a guide). Information on other programs may be found on the UNC System [website](#).
- How this program supports specific university and UNC System [missions](#).
- Collaborative opportunities with other UNC institutions as appropriate.

The proposed Master of Science in Foundations of Data Science is an interdisciplinary program offered by the Departments of Computer Science, Mathematics, and Statistics that will train the next generation of professionals for careers in industry, government, and academia. The program will provide students with advanced skills in the components, methods and tools of data science and their application to a variety of tasks related to knowledge discovery as well as computational and statistical data analysis. The program will not only provide a solid understanding of the foundational concepts of the field but also emphasize collaboration among the field's key disciplines, as advocated by the American Statistical Association, namely database management, statistics and machine learning, as well as distributed and parallel systems. The program is intended to contribute to the economic development of North Carolina by providing a pipeline of experienced data scientists trained to develop data solutions across a range of industries.

A recent study sponsored by the National Science Foundation (http://www.cs.rpi.edu/TFoDS/TFoDS_v5.pdf) concludes that *"theoretical foundations are necessary in all aspects of data science, from the generation and collection of data to the analysis and decision making processes"* and goes on to say that *"particular emphasis should be placed on interdisciplinary collaborations between computer scientists, mathematicians, and statisticians"*. The present proposal addresses these needs head on through innovative graduate training at the cutting edge of this scientific arena.

Professionals completing the program will:

- Design efficient data modeling and processing methods by using mathematical and algorithmic tools.
- Construct conceptual data models, optimize query languages, and implement principles of information integrity, security and confidentiality.
- Quantify appropriate measures of uncertainty associated with the methods of analysis.
- Perform core predictive/descriptive data-mining tasks and design and implement strategies for real-world data-mining problems.
- Develop appropriate data structures and algorithm design techniques including recursion, divide-and-conquer, distributed and parallel optimization, and dynamic programming for analysis of emerging data types.
- Apply statistical learning principles to a variety of data-analysis problems.
- Use relevant software packages and tools and gain insight into how knowledge discovery and data use occurs in practice.

Existing programs in Data Science and/or Analytics provide training in the *usage* of Data Science techniques and applications. The proposed program will fill a void and will instead target the rigorous underpinnings of Data Science, providing a full quantitative perspective into the field. This will enable graduates from this program to gather a more in-depth understanding of not only the usage, but also the development of the methods in data science. To the best of our knowledge, the MS in Statistics: Data Science program at Stanford University is the only other program whose curriculum aligns with our proposed training.

Within the UNC system, the current programs that are most similar to the proposed program are NC State's Master of Science in Analytics (MSA) and UNC Charlotte's Professional Science Master's in Data Science and Business Analytics (DBSA). As is true of a growing number of programs in the US, these two degrees have a strong bias towards business. This is not surprising, as commercialization of analytics tools and techniques has been

primarily driven by business applications. The proposed degree, on the other hand, will develop education capabilities for data science in the sciences and engineering disciplines. The need for such capabilities was strongly emphasized in a 2017 report by the National Academies of Sciences, Engineering and Medicine.¹

With its strong credentials in math, statistics, and computer science, NC State is well positioned to lead the nation in developing unique formal training in data science that covers the key concepts above, including domain-specific considerations. We note that NC State recently created a new PhD in Geospatial Analytics², an interdisciplinary research program that applies analytics to a specific discipline. This request to plan draws from existing NC State expertise to create a complementary educational program in Data Science with a science and engineering focus. We also expect that graduates of the program will be qualified for the Geospatial Analytics program and similar research programs in other science/engineering disciplines.

We believe that the new degree does not represent an unnecessary duplication to NC State’s MSA but rather a complementary educational offering that strengthens our presence in Data Science. Specifically, the new MSFDS program will differ from the MSA in several dimensions, including the educational approach of the new program, its focus and objectives, its content and delivery modalities, and its target audience.

No.	Commitments	Alignment (H, M, L)
1	Academic excellence and the opportunity for success for all students	H
2	Value for students and for North Carolina	H
3	Solutions to North Carolina’s biggest challenges	H
4	Connection and engagement with North Carolina communities	H

The proposed program aligns with the 2013-18 UNC COMPACT priorities (see the above table). More specifically, it advances the broader UNC system directives: Goal 3: “serving the people of NC,” and Goal 4: “maximizing efficiencies” by combining and leveraging existing strengths at NC State in the Statistics, Mathematics and Computer Science programs. In addition, the proposed program strongly aligns with four of the five goals of NC State’s strategic plan (The Pathway to the Future) as noted in the following table.

¹ <http://www.nap.edu/24886>

² <https://cnr.ncsu.edu/geospatial/academics/phd-in-geospatial-analytics/>

NC State – The Pathway to the Future – Strategic Plan 2011 – 2020

No.	Goals	Alignment (H, M, L)
1	Enhance the success of our students through educational innovation	H
2	Enhance scholarship and research by investing in faculty and infrastructure	L
3	Enhance interdisciplinary scholarship to address the grand challenges of society	H
4	Enhance organizational excellence by creating a culture of constant improvement	H
5	Enhance local and global engagement through strategic partnerships	H

There are no identifiable opportunities for collaboration with other UNC institutions. All programs are designed as terminal Masters degrees with a specific curriculum; our focus is on foundations while the others are focused more on applications.

Student Demand: *(Provide evidence of student demand. Discuss the extent to which students will be drawn from a pool of students not previously served by the institution. Maximum length 1,000 words.)*

The DGPs of the Statistics, Mathematics and Computer Science programs regularly receive a sustained number of inquiries from students who would be best served by the proposed degree. While both the Computer Science and Statistics Departments offer a Data Science concentration within their Masters programs, these concentrations only consist of a few elective courses within the discipline. The concentrations, while providing more focus on Data Science than the traditional Masters degree in Computer Science or Statistics, are still single discipline degrees, and do not encompass the truly interdisciplinary nature of Data Science, which can only be met by a program that is developed across disciplines.

We expect that most students will be drawn from a pool of professionals with interests at the intersection of Computer Science, Mathematics, and Statistics. As an example, Computer Science and Statistics currently offer a Data Science Certificate program through NC State Executive Education that is designed for Cisco employees and enrolls approximately 100 students from various Cisco sites around the world every Spring semester. Even if only 10% of these students express interest in a Masters degree in Data Science, that would represent a significant pool of applicants given that these numbers represent a single company in one industry sector. These professionals are not currently served by the University. Moreover, few programs nationwide cater to this pool of students (refer also to Section 5c below).

Societal Demand: *(Provide evidence of societal demand and employability of graduates from each of the following source types. Maximum length 1,000 words)*

- Labor market information (projections, job posting analyses, and wages)
 - o Specific to North Carolina (such as ncworks.gov, nctower.com, or outside vendors such as [Burning Glass](http://burningglass.com))
 - o Available from national occupational and industry projections (such as the [U.S. Bureau of Labor Statistics](http://www.bls.gov))

A Burning Glass analysis shows that among the top 13 states with the greatest demand for the proposed program's graduates, four are in the Southeast (Georgia, Maryland, North Carolina and Virginia). In addition, Washington D.C. is also an area of high demand. The same study demonstrates that the counties of Durham, Orange and Wake account for over 51% of the demand for the proposed program's graduates where the demand is strong and growing. For North Carolina, the projected growth rate for data scientists over the 2016-2026 time frame is 16.69%.

The Bureau of Labor Statistics (BLS) projects above average employment growth for occupations relevant to the proposed program, see https://www.bls.gov/emp/ep_table_102.htm. More precisely, for the 2016-26 period, the BLS projects roughly a 30% employment growth for mathematicians and "software developers, applications" along with a 34% for statisticians.

- Projections from professional associations or industry reports

Glassdoor has named Data Scientist as the best job in America for 2016 (https://www.glassdoor.com/List/Best-Jobs-in-America-LST_KQ0,20.htm) with a median base salary of \$116,840. McKinsey has estimated that the US will face a shortage of 140-190K persons with analytics skills by 2018 (<http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation>).

- Other (alumni surveys, insights from existing programs, etc.)

The Computer Science department offers a Data Science concentration within its Masters program that became official in the 2015-2016 academic year. Eight Fall 2015 graduates, sixteen Spring 2016 graduates, and eighteen Fall 2016 graduates completed this concentration. We expect these numbers to continue to increase during this academic year. The Statistics department has just received approval to begin a Data Science concentration within its Masters program as of Fall 2016. As discussed in Section 3, these concentrations do not capture the full interdisciplinary nature of the foundations of Data Science.

For Doctoral Programs Only:

Describe the following (maximum length 2,000 words):

- The research and scholarly infrastructure in place (including faculty) to support the proposed program.
- Method of financing the proposed new program (including extramural research funding and other sources) and indicate the extent to which additional state funding may be required.
- State the number, amount, and source of proposed graduate student stipends and related tuition benefits that will be required to initiate the program.

Contact: (List the names, titles, e-mail addresses and telephone numbers of the person(s) responsible for planning the proposed program.)

Position Title	Name	E-mail Address	Telephone
Computer Sc. DGP	George Rouskas	rouskas@ncsu.edu	919-515-3860
Mathematics DGP	Pierre Gremaud	gremaud@ncsu.edu	919-515-3085
Statistics DGP	Wenbin Lu	wenbin_lu@ncsu.edu	919-5151915

This Letter of Intent to Plan a New Program has been reviewed and approved by the appropriate campus authorities.

Position Title	Signature	Date
Provost		
Provost (Joint Partner Campus)		

Master of Science in Data Science

SUMMARY OF ESTIMATED ADDITIONAL COSTS FOR PROPOSED PROGRAM

INSTITUTION	North Carolina State University	DATE	25-Feb-17
Program (CIP, Name, Level)	30.3001		
Degree(s) to be Granted	Master of Science	Program Year	Year 1 (2019-2020) Year 4
Differential tuition requested per student per academic yr	\$5,600		
Projected annual FTE students		23	46
Projected annual differential tuition		\$112,000	\$224,000
Percent differential tuition for financial aid		15%	15%
Differential tuition remainder		95200	190400

ADDITIONAL FUNDS REQUIRED - BY SOURCE

	Reallocation of Present Institutional Resources	Projected Differential Tuition	Enrollment Increase Funds	Other New Allocations (Identify)	Total
EPA/SPA Regular Salaries					
graduate service coordinator	\$	33,250.00	\$ -	\$ -	\$ 33,250.00
EPA Academic Salaries					
director of graduate program	\$	55,417.00	\$ -	\$ -	\$ 55,417.00
faculty fellow	\$	20,000.00			\$ 20,000.00
Social Security (fringe included above)	\$ -	\$ -	\$ -	\$ -	\$ -
State Retirement	\$ -	\$ -	\$ -	\$ -	\$ -
Medical Insurance	\$ -	\$ -	\$ -	\$ -	\$ -
Graduate Stipends					
(3 Tas @ \$22K+fringe)	\$	76,560.00	\$ -	\$ -	\$ 76,560.00
Supplies and Materials					
office supplies and computer for GSC	\$	2,500.00	\$ -	\$ -	\$ 2,500.00
Current Services	\$ -				
(Identify)	\$ -	\$ -	\$ -	\$ -	\$ -
Travel	\$ -	\$ -	\$ -	\$ -	\$ -
Communications	\$ -	\$ 500.00	\$ -	\$ -	\$ 500.00
Printing and Binding	\$ -	\$ 500.00	\$ -	\$ -	\$ 500.00
Advertising	\$ -	\$ 1,000.00	\$ -	\$ -	\$ 1,000.00
Fixed Charges	\$ -				
(Identify)	\$ -	\$ -	\$ -	\$ -	\$ -
Capital Outlay (Equipment)	\$ -				
(Identify)	\$ -	\$ -	\$ -	\$ -	\$ -
Libraries	\$ -	\$ -	\$ -	\$ -	\$ -
TOTAL ADDITIONAL COSTS		\$ 189,727.00	\$ -	\$ -	\$ 189,727.00
	\$ -				

Narrative: The program will be managed by a Director who will be compensated through 1 Summer month of salary and a yearly 1 course buyout. The Director will be assisted by a graduate service coordinator (.5 FTE).

The teaching will be supported by three teaching assistantships (\$22K each), one in each of CS, Math and Stat. In addition, the program will have available a \$20K for buyout available to CS, Math and Stat faculty interested in developing or updating courses relevant to the Program.

It is expected that the program will ramp up to a size of about 50 by year 4 at which time it will be entirely self-supported.