Proposal for New Program:
M.S. in Data Science: Computational Analytics

1. “Rationale for the new program...” The proposed “Data Science: Computational Analytics” M.S. is designed for students interested in developing expertise in data science, with specialization in computational analytics. Data Science is an interdisciplinary discipline about methods and systems to extract knowledge or insights from large quantities of data coming in various forms. Data science employs techniques and theories drawn from many fields within the broad areas of mathematics, statistics, and computer and information sciences and applies them on a wide range of data-rich domains such as biomedical sciences, physical science, geoscience, social science, engineering, business, and education.

Data science is a very broad and multifaceted field. The “Computational Analytics” emphasis will enable students to analyze large quantities of data to discover new knowledge and facilitate decision making. To accomplish this, the programs will provide students with a strong foundation in big data management and analysis, the algorithmic, computational, and statistical thinking, and understanding of computer systems.

Graduates of this program will have the necessary skills to find data science-related jobs which require a mix of data analysis skills, ability to deal with large quantities of data, and a strong foundation in computer science. We expect that a subset of students will choose to continue with doctoral studies, either to deepen their overall data science expertise or learn how to better use their analytics skills in a particular data-rich domain.

a. “Regional and national competitors...” Following the rapid growth of the field, there is an increasing number of undergraduate and graduate programs in data science. The following web site is attempting to keep track of the graduate-level programs related to data science: [http://www.mastersindatascience.org/schools/](http://www.mastersindatascience.org/schools/). It could be observed that there are roughly 5 homes for those programs: (1) computer science departments (with emphasis on computational analytics), (2) schools of business (with emphasis on business analytics), (3) departments of statistics (with emphasis on statistical data analysis), (4) domain-specific departments (many having the “informatics” in the name, such as bioinformatics, healthcare informatics, social informatics), (5) multi-departmental programs (focusing on a holistic approach to data science).

Regionally, most of the existing programs have a focus on business analytics. In Fall 2015 Wharton School at Penn started offering online “Business Analytics” specialization at Coursera ([https://news.wharton.upenn.edu/press-releases/2015/09/wharton-school-offer-new-business-analytics-specialization-coursera-online-learning-platform/](https://news.wharton.upenn.edu/press-releases/2015/09/wharton-school-offer-new-business-analytics-specialization-coursera-online-learning-platform/)), whose focus is on using data science in business domain. Villanova started offering an M.S. in Analytics that focuses on business analytics ([https://www1.villanova.edu/villanova/business/graduate/specializedprograms/msa.html](https://www1.villanova.edu/villanova/business/graduate/specializedprograms/msa.html)), as well as Master of Science in Applied Statistics with emphasis on statistical data analysis, Drexel is offering an M.S. in Business Analytics ([http://www.lebow.drexel.edu/academics/graduate/masters/ms-business-analytics](http://www.lebow.drexel.edu/academics/graduate/masters/ms-business-analytics)), as well as Online Master's in Health Informatics with emphasis on health data, and Saint Joseph’s has
Online Master of Science in Business Intelligence and Analytics. The proposed program will be the first data science program in the region with the emphasis on computational analytics.

Nationally, there is an increasing number of programs with similar emphasis on computational analytics, such as those offered by computer science departments at Carnegie Mellon (Master of Computational Data Science), Columbia (Master in Data Science, MS in CS concentration in Machine Learning), NYU (MS in Data Science), and many others.

b. “Disciplinary trends…” Data science is a rapidly growing field with an increasing demand in industry, research, and government. A recent McKinsey Global Institute study states that the US will face a shortage of about 190,000 data scientists and 1.5 million managers and analysts who can understand and make decisions using big data by 2018. Seventy four percent of the respondents of the recent ‘Peer-Research Big Data Analytics Survey’ have agreed that data science is adding value to their organization and allows vital information for making timely and effective business decisions of great importance. In a recent MIT Sloan Management Review survey, four in ten (43%) companies report their lack of appropriate analytical skills as a key challenge.

c. “Comparison to top programs…” CIS Department is well positioned to offer a strong program for three main reasons: (1) the program design is consistent with the existing graduate data science programs offered by top-ranked computer science departments, (2) CIS Department has a strong core of research-active data science faculty that will make sure the program will be and remain current, (3) most of the courses in the program already exist and they are designed and taught by experts in computational analytics and computer science. Those factors will result in a vibrant program which will attract talented and ambitious students and teach them the tools to have successful careers in data science.

2. “Relationship to other programs at Temple…” This program is the most similar to the M.S. in Computer Science. The main difference is that, by focusing on data science and computational analytics, the program is less broad and less demanding in traditional computer science (less emphasis on theory and computer systems) and stronger and broader in machine learning, data analysis, and management of big data. As a result, the two programs have different core and elective course requirements. See Appendix C for more detail. There is some similarity with Ph.D. in Statistics, because both programs focus on data analysis. The proposed program is much more focused on the computational aspects of data analysis, which require significant background in computer science. It is worth noting that the proposed program counts many of the graduate statistics courses as electives.

3. “Curriculum…” The summary of the program is provided in Appendix A and a sample 4 semester grid is provided in Appendix B.

a. The program requires 30 credits. It has 4 core courses aimed at giving the strong background in programming, computer systems, algorithmic thinking, and machine learning. There are 5 electives from roughly 4 groups of courses: data analytics, dealing with big data, statistics, and applications. Students will not be able to register only for electives in one of those groups. Students are also allowed to register for up to 3 credits of independent studies to gain more practical experience by working directly with
research-active data science faculty at CIS Department. Each student is expected to do
the M.S. Project in Data Science (3 credits).
b. This is a 30 credit course that will allow student to finish it within 2 years for full-time
students and within 2.5 years for part-time students.
c. The program is only offered on Main campus.

4. “Impact on Faculty & Students...”

a. All core and elective courses already exist. The listed core and elective CIS courses are
already being regularly offered as part of the existing M.S. in CS and Ph.D. in CIS
programs in the department. The faculty who are and will be teaching these courses are
already in place. If the number of students admitted to the program is controlled, there
will be no need to open new sections of any course, but it is likely that we will need to
provide teaching assistant support to the high demand courses. Only in the event the
interest of students in the program exceeds our expectations will the CIS Department
have to reconsider whether some of the courses will need to be offered more frequently
and be taught in larger classrooms and how will this be handled.
b. The admission requirements will be consistent with other M.S. programs at Temple. The
program will be advertised together with other graduate programs at CIS through our
website as well as the university website and print materials. We are expecting a
reasonable demand for this program and anticipate plentiful job opportunities (see answer
1.b). The new program will complement our existing M.S. in CS program by attracting
applicants with relatively stronger mathematical and thinner CS backgrounds (such as
those with degrees in engineering and sciences).

5. “Impact on Resources...”

a. There is no special tuition for this program above the regular graduate tuition.
b. Students are expected to be self-supported. A very limited number of teaching and
research assistantships might be available for the students from this program.
c. We project 10 new students per year, with the objective of filling our currently offered
graduate courses to capacity. Assuming that in average 24 out of 30 credits will be from
courses offered by CIS, the program is expected to contribute 240 credit hours each year.
d. Since each student is projected to take 6 credits outside of CIS, we project a slight
increase in demand for some of the non-CIS elective courses. Given a large array of
possible electives, the increased demand will not put an excessive burden on any
particular course.
e. If the number of new students is constrained at 10 per year, there will be very little
impact on space resources. However, if the demand becomes larger than projected, CIS
and CST will have to decide whether the increased enrollment could be handled in a cost-
effective way.
f. 1. There will be a modest increase in tuition revenue and costs due to the increase in
students of 10 per year, but the programs will continue to make Temple an attractive
place for study.
   2. There might be a need to hire several teaching assistants to support the instruction
6. “**Assessment**...” The program goals are listed in Appendix D. Assessment will continue as it has before through the Assessment committee of the department.

7. “**Summary of peer and aspirant programs**” See discussion in part 1.

8. “**Implementation**...”
   
a. We would like this program change to begin in Fall 2016.
   b. Continuing students in related programs will be able to continue in their current programs or change to the new program.

9. “**Process**...” Dean Michael Klein and Associate Deans Shohreh Amini and Robert Levis were consulted. The CST Graduate Committee was consulted and unanimously approved the proposal. The CST Dean’s Advisory Committee was consulted and unanimously approved the proposal. The CST Collegial Assembly was consulted and unanimously approved the proposal.
Appendix A
M.S. in Data Science: Computational Analytics: Details

(30 credit hours)
(Core Courses: 4 courses)
- Programming Techniques (CIS 5511)
- Pick one from (Operating Systems (CIS 5512), Principles of Data Management (CIS 5516), Computer Architecture (CIS 5642))
- Machine Learning (CIS 5526)
- Design and Analysis of Algorithms (CIS 9615 – renumber to 5615)

(Electives: select 5 from)
Select at most 3 from Data Analysis electives:
- Data Mining (current CIS 9664, renumber to 5523)
- Analysis and Modeling of Social and Information Systems (CIS 5524)
- Neural Computation (CIS 5525)
- Probabilistic Graph Models (CIS 5535)
- Text Mining and Language Processing (CIS 5538)
- Computer Vision (CIS 5543)
- Artificial Intelligence (CIS 5603)

Select at most 3 from Big Data electives:
- Principles of Data Management (CIS 5516)
- Advanced Topics in Data Base Systems (CIS 9665)
- Operating Systems (CIS 5512)
- Computer Architecture (CIS 5642)
- Distributed Computing (CIS 5644)
- Emerging Storage Technologies (CIS 5xxx)
- Data-Intensive and Cloud Computing (a new cross-listed course: 5xxx)
- Topics in Computer Science (CIS 5590) (subject to approval from advisor)

Select at most 3 credits of Independent Study
- Independent Study (subject to approval from advisor)

Select at most 1 CIS course with number less than 55xx (subject to approval from advisor)
Select at most 2 statistics and domain-related courses (subject to approval from advisor).
Some examples:
- From M.S. in Statistics (any 8000- or 9000-level STAT course)
- From M.S. in Biology (BIOL 5312 Biostatistics, BIOL 5225 Evolutionary Genetics Genomics, BIOL 5229 Systems Biology: Principles and Applications, BIOL 5403 Genomics, BIOL 5411 Structural Bioinformatics, BIOL 5509 Computational Genomics, BIOL 5511 Ethics in Bioinformatics)
- From M.S. in Geology (EES 5011 Remote Sensing and GIS, EES 8421 Groundwater Modeling)
- From M.S. in Mathematics (MATH 5063 Introduction to High-Performance Computing Technology for Scientists, MATH 8013 Numerical Linear Algebra, MATH 8031 Probability Theory, MATH 8007 Introduction to Methods in Applied Mathematics I)

**M.S. Project 3 credits**

- Project in Data Science *(CIS 9992, this is new course similar to CIS 9991)*
## Appendix A
### M.S. in Data Science: Computational Analytics: An example sequence

<table>
<thead>
<tr>
<th>Fall – year 1 (9 cr.)</th>
<th>Spring – year 1 (9 cr.)</th>
</tr>
</thead>
</table>
| CIS 5511 Programming Techniques  
CIS 5526 Machine Learning  
CIS 5512 or CIS 5516 or CIS 5642 | CIS 5615 Design and Analysis of Algorithms  
Elective 1 (Data Analytics)  
Elective 2 (Big Data) |

<table>
<thead>
<tr>
<th>Fall – year 2 (6 cr.)</th>
<th>Spring (6 cr.)</th>
</tr>
</thead>
</table>
| Elective 3 (Statistics/Domain-Related)  
Elective 4 (Data Analytics) | Elective 5 (Big Data)  
CIS 9992 Projects in Data Science |
Appendix C
Comparison with related programs

Comparison versus MS in Computer Science

MS in Computer Science (30 credits):

4 Core Courses

- CIS 5511 Programming Techniques 3
- CIS 5512 Operating Systems 3
- CIS 5513 Automata and Formal Languages 3
- CIS 9615 Design and Analysis of Algorithms 3

Plus:

- Project or Thesis Courses 1
  - CIS 9991 Project in Computer Science (for 3 credits) 18
  - CIS 9996 Thesis in Computer Science (for 6 credits) and four electives

Select electives from the following CIS courses, which are grouped topically, or take courses outside the department with the approval of the CIS Graduate Committee: 2

Main Differences:

- CIS 5513 is replaced with CIS 5525 in core requirements: provides the introduction to main concepts in computational analytics
- Instead CIS 5512, there is a menu choice of CIS 5512, 5516, 5642: CIS 5512 is a hard core CS course that requires significant background in computer systems; many non-CS students find this course extremely difficult; CIS 5516 is much more accessible to non-CS students.
- Instead of free choice of 5 electives, there is a carefully crafted set of electives that allow student to improve their breadth and depth in data science while also allowing customization to the specific student interests.
Appendix D

Data Science: Computational Analytics MS Program Goals

Program Goals

*After completing this program, students should:*

- have the ability to process and analyze effectively real-life data from a wide range of domains.
- have the ability to work with large quantities of data.
- be proficient in applying mathematics concepts from calculus, linear algebra, and statistics to problem solving
- have effective programming, program debugging, and program testing skills
- have the ability to use computational and algorithmic thinking to develop computationally-efficient data science algorithms
- be proficient in using computer technology and software in solving real-life data analysis problems.
- handle unfamiliar concepts and situations and apply disciplined thinking techniques to new settings
- be able to communicate using oral, written, or electronic media, and have the teamwork and leadership skills needed to recognize, isolate, and solve data science problems.
- be committed and open to life-long learning, new ideas, and be able to bring them to bear to help others.