# MARYLAND HIGHER EDUCATION COMMMISSION ACADEMIC PROGRAM PROPOSAL

1	PROPOSAL FOR:		
X NEW INSTRUC	CTIONAL PROGRAM		
	L EXPANSION/MAJOR MOD	IFICATION	
· · · · · · · · · · · · · · · · · · ·	E DEGREE PROGRAM		
WITHIN EXIS	ΓING RESOURCES or X R	EQUIRING NEW RESOURCES	
	<u>separate</u> cover page. For examp a degree program and a certific	ole, two cover pages would accompany a ate program.)	
	Mount St. Mary's University		
	Institution Submitting Proposal		
	Spring 2019		
	Projected Implementation Date		
Post Baccalaureate Certificate  Award to be Offered	Data Scie	pposed Program	
		27.0599	
Suggested HEGIS Code		Suggested CIP Code	
Math and Computer Science		Dr. Melanie Butler	
Department of Proposed Prog	ram — 1	Name of Department Head	
Dr. Boyd Creasman, Provost	b.creasman@msmary.edu	301-447-5218	
Contact Name Contact E-Mail Address Contact Phone Number			
Signature and Date	1/24:/2018 President/Chief	Executive Approval  Approved by Governing Board	
Doto	Dutc Engoised/	approved of Governing Bound	



January 24, 2018

OFFICE OF THE PROVOST 16300 OLD EMMITSBURG ROAD EMMITSBURG, MARYLAND 21727

301.447.5218 FAX: 301.447.5863

Dr. James Fielder, Secretary Maryland Higher Education Commission 6 North Liberty Street, 10th floor Baltimore, MD 21201

Dear Dr. Fielder,

I am writing to submit the enclosed proposal for a Post-Baccalaureate Certificate (PBC) program in Data Science to be offered at Mount St. Mary's University.

The PBC is designed to build data science skills for a broad target audience crossing multiple discipline areas. Data science is relevant across a spectrum of subject areas, and individuals within any discipline would benefit professionally from developing some level of data science expertise. While data science certification and training is prevalent in online forums, prerequisite computing and statistics skillsets are typically intensive, thus discouraging or excluding a large audience with limited experience in either. Bridging this gap for these individuals is the focus of the PBC. Those who complete the certificate will have gained critical data science expertise that can be utilized immediately within their current career paths. Those who choose to continue on with more intensive data science education will have the prerequisite skillsets needed to now participate in the advanced data science curriculum typically offered online.

The proposed Data Science certificate consists of five courses (15 credit hours). It is designed for individuals employed across all professions who are seeking knowledge and career advancement potential related to data science activities. The focus is an augmentation to their previous educational and career experiences as specific subject area knowledge is highly desired within the data science community. While all participants who successfully complete the program will be prepared to pursue more advanced training in data science at the graduate level, all will also be well-positioned to leverage data science skills within their current career paths.

The proposal details the requirements and specifics of the program. Thank you for your time and consideration.

Sincerely,

Boyd Creasman, Ph.D.

Provost



# Mount St. Mary's University

# Proposal for a New Post-Baccalaureate Certificate Program

# **Data Science**

Developed by the Department of Mathematics and Computer Science

# A. Centrality to Institutional Mission and Planning Priorities

#### **Program Description**

A Post-Baccalaureate Certificate (PBC) in Data Science is proposed within the Department of Mathematics and Computer Science at Mount St. Mary's University. The PBC is designed to build data science skills for a broad target audience crossing multiple discipline areas. Data science is relevant across a spectrum of subject areas, and individuals within any discipline would benefit professionally from developing some level of data science expertise. While data science certification and training is prevalent in online forums, prerequisite computing and statistics skillsets are typically intensive, thus discouraging or excluding a large audience with limited experience in either. Bridging this gap for these individuals is the focus of the PBC. Those who complete the certificate will have gained critical data science expertise that can be utilized immediately within their current career paths. Those who choose to continue on with more intensive data science education will have the prerequisite skillsets needed to now participate in the advanced data science curriculum typically offered online.

For the second year in a row in April 2017, <u>Glassdoor</u> reported data scientists in the top spot of the 50 best jobs in America. In a March 2017 article, <u>Inside Higher Ed</u> emphasized "a shortage of job candidates with fluency in data science and analytics is among the nation's most yawning of skills gaps." The article stated a need for fixes by both higher education and businesses, and a recommendation that all undergraduates should complete coursework in data science. With the explosion of data availability across industry, service, government and research organizations, the demand for data scientists is equally wide spread across a diverse discipline spectrum. The need for individuals who have knowledge of their subject area and knowledge of the means to leverage the vast quantity of data related to that subject area to support decision-making has become critical.

According to the <u>2016 US Census</u>, Frederick County, Maryland is among the fastest growing suburbs of the Baltimore-Washington metropolitan area, projected to grow 38% beyond current census by the year 2040. Employers within the county include government, research, and financial spectrums as well as many startup high technology companies including biotechnology and cybersecurity. Additionally, many of the residents of the county are employed across an equally broad spectrum of careers within the larger Baltimore-Washington metropolitan area.

The proposed Data Science certificate is designed for individuals employed across all professions who are seeking knowledge and career advancement potential related to data science activities. The focus is an augmentation to their previous educational and career experiences as specific subject area knowledge is highly desired within the data science community. While all participants who successfully complete the program will be prepared to pursue more advanced training in data science at the graduate level, all will also be well-positioned to leverage data science skills within their current career paths. This certificate will be equally beneficial for those who do not intend to pursue data science as a career direction, but simply desire to be knowledgeable related to data and its potential utilization within their organization.

The full PBC program curriculum consists of five courses (15 credits) offered through a combination of MSMU's Emmitsburg and Frederick campus locations and integrating online elements throughout. The three foundational courses will be offered as hybrid courses. Cohort sessions will be scheduled to use one campus (Frederick or Emmitsburg) for that cohort, and will leverage the same facility for all subsequent in-person instruction sessions for that course. Utilizing MSMU's Frederick and Emmitsburg campuses for alternating cohorts will enable the program to reach a geographically disparate population. Initial course sessions for each cohort will be on campus with the instructor to establish connection across cohort participants and to establish comfort with technology platforms leveraged throughout the curriculum. Subsequent versions of the foundational courses will be instructor-led through online video technology with periodic sessions face-to-face at the cohort site location. The applications course will be offered online as a cross-disciplinary effort across the MSMU curriculum. The final course is a capstone project selected by the participant with oversight and final approval from a cross-disciplinary panel of faculty. Students attending part-time will be able to complete the certificate in one year.

The design of the program ensures participants develop key foundational computational skills in statistics, data engineering, and data model applications prior to going deeper within subject areas. No assumptions will exist regarding the participants previous experiences in either technology, programming, or statistics. The foundational courses within the certificate curriculum focus on the many formats, interfaces, and means by which data can be acquired as well as on building high-demand technical skills related to the data science lifecycle and "big data." For participants with some previous experience in programming and statistics, there will remain substantial new material within these foundational courses to augment their previous knowledge.

This PBC will build on common foundational courses in statistics and data engineering in existence within MSMU's Data Science minor curriculum. Because the target audience is cross-disciplinary, the curriculum will heavily leverage interdisciplinary applications of data science. Mount St. Mary's University is well-suited to providing cross-disciplinary flavor to data science

as numerous professors are actively leveraging aspects of data science within their research and classroom experiences. Current application areas include biostatistics, psychology, criminal justice, cybersecurity, and political science among many others. MSMU is also well-suited to providing data science instruction within an ethics-focused framework that ensures participants never lose sight of the potential for individual impacts related to data science technology decisions.

Within the PBC curriculum, the first three courses are designed to provide a critical, common foundation in key computing and statistical data science concepts. An application course then demonstrates real-world utilization of data science skills and technology across a spectrum of four or five application areas. For the first few iterations of the PBC, the application course will include a set list of topics. This commonality ensures seamless integration of application curriculum with the specific technology skills obtained by participants through the foundational courses. Over time, new topic areas will continue to be developed and integrated within the application platform to ultimately allow participants selection of their own combination of four or five application topics. Once fully integrated across MSMU curriculum, participants will have sufficient variety of topic areas available to complete certificates within specific tracks such as bioinformatics, cheminformatics, DevOps informatics, and cybersecurity informatics or to create a personalized experience that crosses a variety of disciplines of interest. The capstone course in the certificate sequence provides participants the opportunity to hone their skills within a specific topic area reviewed and approved by a cross-disciplinary panel of university faculty.

Creating a new curriculum within each subject area becomes problematic as the level of overlap related to statistics and data engineering can be substantial. The proposed program takes this overlap into account by providing individuals who have already completed a bachelor's degree within specific subject areas the opportunity to learn common foundational skills and then demonstrate those skills with mentoring and oversight from a combination of data science faculty members and faculty members aligned with their specific discipline.

#### Relationship of Proposed Program to the University's Mission

The proposed program is designed to be consistent with and to support the University's mission:

Mount St. Mary's is a Catholic university committed to education in the service of truth; we seek to cultivate a community of learners formed by faith, engaged in discovery, and empowered for leadership in the Church, the professions, and the world.

As a testament to the university mission, MSMU has a long-standing history of offering curriculum, including technology curriculum, within a structure of smaller, student-focused, class sizes, fully integrated within an ethics-focused framework. The proposed certificate program builds on this success in order to empower participants to develop skills and insights necessary to be ethical data leaders within their individual professions.

Basing the foundational courses on existing data science courses enables the program to inherit the framework already in place. The applications course will be created through a montage of curriculum elements already in place from several existing data applications courses. Basing the PBC curriculum on proven material ensures alignment with the University's mission statement specific to formation of students. Participants of the certificate program will develop as data literate professionals well-prepared to function analytically, professionally, and ethically within a high-technology world.

As data science solutions continue to shape the world around us, it is imperative that universities like Mount St. Mary's University participate in the preparation of professionals who "see and seek to resolve the problems facing humanity, and commit themselves to live as responsible citizens." Participants who successfully complete the certificate program will think intellectually as well as morally related to the utilization of data and associated data models. The program will develop not only the technical skills specific to data science, but will challenge participants to never lose sight of the ethical implications of future data science products.

### Relationship to Strategic Goals

As a university community, Mount St. Mary's has been participating in a comprehensive effort to refine our mission and strategic plan. The initial revisions are on track for completion by the fall of 2017. As part of these activities, the President in conjunction with the full university community developed shared priorities for AY 2016-17. The following are excerpts from those shared priorities:

- 1. Improve the excellence of our educational experience by engaging the entire Mount community.
  - a. Continue to improve the integration of curricular and co-curricular programs.
  - b. Increase online learning courses available at both the Emmitsburg and Frederick campuses
- 2. Prioritize and develop the most promising areas for growth consistent with our core values as a Catholic institution of higher education.
  - a. Research and develop new undergraduate and graduate programs that meet the current and future needs of students and employers

In support of these efforts, a committee was established under direction of the Provost's Office in the fall of 2016 to identify potential new academic programs for consideration. That committee composed of administration, faculty, and student representatives reviewed numerous proposals from faculty and selected a few representative of desired future direction for the university. The proposed data science certificate was one of the proposals selected by the review committee as providing value in alignment with the university direction.

Additionally, as part of the larger comprehensive effort related to mission and strategic plan, insights for relevant new programs were sought from Board members specific to their own industry experiences. Data and informatics were a common theme in those discussions. The certificate program as proposed would be designed to provide common technology elements while also being adaptable to the changing informatics needs across a broad spectrum of disciplines.

# B. Adequacy of Curriculum Design and Delivery

### **Program Requirements**

#### **Purpose**

The certificate in Data Science is designed for professionals from diverse backgrounds and disciplines with a goal to build on the subject matter expertise within the individual applicant's professional and educational experiences. It is designed to provide participants with the knowledge and skills necessary to contribute at an expanded level within a highly data-centric world. Participants will develop skills within key technology areas, key computational areas, and key application areas without presumed prior knowledge specific to these areas.

#### Admission Requirements

Candidates for admission into the data science certificate program must have completed a bachelor's degree and must satisfy the following criteria:

- 2.75 minimum cumulative undergraduate or graduate grade point average (GPA)
- International students must also submit a TOEFL score with their application.

#### Course List (5 courses, 15 credits)

#### Required Courses (5 courses, 15 credits)

DATA 5XX: Introduction to Data Science and the Data Science Lifecycle (3 cr.) — This course presents an overview of the discipline of data science: its goals, methods, tools, and scope. Additionally, the course provides understanding of the end to end lifecycle of a data science initiative, and the variety of data products in today's world that are a result of applied data science. As one of the top skillsets sought within the data science community, the R statistical computing environment is leveraged throughout the course. Projects throughout course provide opportunity to demonstrate mastery of course concepts specific to data manipulation, statistical analysis, and associated exploratory visualizations against a "tidy" data set. Ethical issues surrounding data collection and use are discussed.

**DATA 5XX: Data Engineering and Data Munging (3 cr.)** – This course focuses on the data acquisition, cleaning, manipulation, transformation, and analysis portions of the data science lifecycle. Data engineering and data munging techniques related to a variety of formats are covered. Open data repositories from government organizations as well as social media sites and databases are explored. The R computing environment continues to be leveraged while also investigating numerous ancillary tools within the R community related to data transformation and visualization. Concepts that were discussed at a higher level in the introductory course are now investigated more fully against real world raw

data that requires manipulation to achieve the "tidy" format. Projects throughout course provide opportunity to demonstrate mastery of data science concepts from data acquisition through data engineering and statistical analysis portions of lifecycle.

Prerequisite DATA 5XX: Introduction to Data Science and the Data Science Lifecycle

DATA 5XX: Predictive Modeling Algorithms and Data Products (3 cr.) – This course focuses primarily on the data product portion of the data science lifecycle, and the associated modeling algorithms. Topics are selected from statistical inference, regression, classification, machine learning, natural language processing, and neural network models. The course seeks to examine the assumptions, capabilities, limitations, and advantages of these models within the context of application areas. The R computing environment continues to be leveraged along with expansion to other key skills such as advanced data visualization including higher dimensional data. Projects throughout course provide opportunity to demonstrate mastery of data model and data product concepts. Prerequisite DATA 5XX: Data Engineering and Data Munging

DATA 5XX: Data Science Applications (3 cr.) – This course explores a variety of enterprise level data science applications across a broad spectrum of disciplines. While earlier data science courses focused on distinct portions of the data science lifecycle, investigation within this course follows the full end-end lifecycle for each selected application. Mastery of previous data science concepts enables students to now fully question and understand decisions made throughout the data science lifecycle. Topics selected rely on the analytic and technology skills from earlier courses, specifically R and associated tools. The commonality of data science lifecycle, approach, and technology, regardless of subject area, is emphasized. Students complete a high level capstone project proposal in preparation for Data Science Capstone. *Prerequisite DATA 5XX: Predictive Modeling Algorithms and Data Products* 

**DATA 5XX: Data Science Capstone (3 cr.)** – This course is a culmination of the Data Science sequence enabling the student to demonstrate skills learned within a project area of interest for their future or current direction. Proposals for projects are submitted and reviewed by a cross-disciplinary faculty panel prior to project initiation. Final project completion is also reviewed and approved by a cross-disciplinary faculty panel for successful completion. *Prerequisite DATA 5XX: Data Science Applications* 

#### **Educational Objectives and Student Learning Outcomes**

The PBC in Data Science is designed to provide key technology and data skills for professionals from all professional areas to be confident in decision making tied to their subject area data. Completion of the certificate will provide students with:

- A detailed understanding of the Data Science lifecycle
- Working knowledge of key technology skills commonly utilized within data science across a broad spectrum of disciplines
- Detailed knowledge of the application of data science models within a variety of disciplinary backgrounds

- Hands on experience in building data science product from a variety of data sources
- An ethical foundation to be leveraged in data decision-making

### Discuss how general education requirements will be met, if applicable.

Not applicable.

#### Specialized accreditation or graduate certification requirements

Not applicable

#### Contracting with another institution or organization

Not applicable.

# C. Critical and compelling need as identified in the 2013 State Plan

The proposed certificate is aligned with many of the goals stated in the <u>Maryland Ready 2013-2017 Maryland State Plan for Post-Secondary Education</u>. Specific goals are identified and aligned with the certificate program objectives below.

# Quality and Effectiveness Goal:

As part of the overall curriculum of Mount St. Mary's University, the certificate program will garner the same benefits as the overall university curriculum in terms of assessment processes and faculty professional development funding. The program will be continuously assessed and evaluated against formal plans to ensure the courses and program align with learning goals and objectives, and with the mission of the university. Those results will be used to provide constant improvement to the material and delivery mechanisms leveraged.

The faculty focused on delivering courses within this certificate are full-time members of the faculty, have significant experience teaching in the subject areas, and several bring industry experiences to their students as well.

With a combination of hybrid and online delivery, this program will follow the successful approaches of similar delivery across the Mount curriculum. Mount St. Mary's has made a concerted effort to collaborate across the faculty related to successful online and hybrid delivery practices and approaches. A faculty committee is in place to oversee those guidelines in collaboration with School of Education resources.

Access, Affordability and Completion Goal:

The primary participants in this certificate program are anticipated to be adult learners who are attempting to augment their current career path with relevant technology skills. The design of the courses with a hybrid and online focus is intended to minimize impact to their busy lives by providing a flexible delivery mechanism to the maximum extent possible. This follows the success models across current MSMU graduate programs related to evening courses and the flexibility of online/hybrid delivery.

Additionally, it is understood that many of the certificate participants will come into the program with minimal advanced experience in mathematics and computer science. No assumptions related to readiness within these subject areas is assumed, which deviates from other programs regionally as well as many online curriculum. The program is designed to build these skills for all participants regardless of previous experiences.

Exposure to current technology within data science will be an integral part of the full educational experience within the program. As the industry relies heavily on open source technology, the intent of the program will be to leverage comparable open source technology as much as possible. This will enable participants to gain experience with specific toolsets while also keeping participant costs manageable.

Affordability is a key element for successful recruitment of certificate participants. The opportunity to compare with comparable programs in the state of Maryland is limited as Data Science certificate programs tend to be focused on graduate level training with computer science and mathematics pre-requisites. As a basis of comparison, the per credit hour for data science certificate programs, if they exist, as well as computer science master's programs have been included in the table below. Additionally, the total cost for certificate completion has been included.

Inst.	Graduate Program	\$ Per Gredik Hours	S Per Courte. (5 of)	Cost of Completed Confilings
MSMU	Science – Graduate Level	\$635	\$1905	5 courses, total \$9525
Loyola University of Maryland	M.S. Data Science	\$1000	\$3000	10 courses, total \$30,000
Johns Hopkins	PMC Data Science M.S. Data Science	Per course tuition	\$4055	MS 10 courses, total \$40,550 PMC 6 courses, total \$24,330
UM College Park	Certificate Data Science	\$458 for Maryland residents, \$659 for out of state residents.	\$1374 in-state \$1977 out of state	12 credits, total \$5496 instate and \$7908 out of state
Towson University	M.S. Computer Science	\$379 for Maryland residents, \$785 for out of state residents	\$1137 in-state \$2355 out of state	N/A
UMBC	M.S. Computer Science	\$621 for Maryland residents \$1047 for out of state residents	\$1863 in-state \$3141 out of state	N/A
Hood	M.S. Computer Science	\$500	\$1500	N/A

#### • Innovation Goal:

The strength of this proposed certificate is its innovative nature. Traditionally, data science skills are taught at the graduate level to individuals who have significant experience in computational technology programs such as computer science, mathematics, and statistics. The target participants for this program are not this group of high-technology professionals, but rather individuals who for any number of reasons may not have significant experience with computational technology. For these individuals, the gap to be bridged to anticipate future careers that heavily rely on this technology can be daunting. The program is designed to provide them a first bridge in crossing that gap.

Additionally, leveraging a combination of online and hybrid structure in conjunction with current faculty and course curriculum ensures delivery of a proven framework in a format desirable for our target participants. Online certificates in data science can be found through universities throughout the United States; online education has become more mainstream. Many of these programs however have fairly significant prerequisite requirements related to programming, statistics, and mathematical maturity levels. Participants in these programs often have fairly advanced technology skills as well. The innovative approach with this program is to leverage instructor-led online elements in conjunction with classroom experiences that enable individuals less comfortable with computational technology to be fully successful.

The technology skills that will be taught are highly desirable skills in the marketplace today related to data and data science. Technology tools and skills change on a rapid basis. Part of the challenge and focus of this curriculum will be to provide participants with resources beyond tools alone that empower them to stay in tune with the changing technology, to find the locations for problem resolution as they begin to utilize the skills learned, and to learn to leverage many online resources that may be unfamiliar prior to enrollment in the curriculum. Participants will learn how to continue learning within the subject area by continually leveraging online resources integrated throughout the course curriculum.

#### • Economic Growth and Vitality Goal:

As highlighted in the State Plan,

"...students...must have access to high-caliber and effective training that meets the evolving needs of the workplace..."

"Science, technology, engineering, and mathematics (STEM) occupations have been identified as an area of high need in Maryland..."

The Mathematics and Computer Science department and the School of Natural Science and Mathematics at MSMU have a long-standing history of placement of graduates

with top STEM employers in the region. Within the last 3 years, MSMU Mathematics and Computer Science graduates have been employed by Northrup Grumman, Raytheon, Lockheed Martin, Johns Hopkins Advanced Physics Laboratory, Patuxent River Naval Air Station, NSA, SAIC, and Booz Allen Hamilton among many others. The proposed certification program will rely on existing advisory connections with regional STEM employers. Basing the foundational courses on the existing Data Science minor required courses ensures that the same high demand data science technology skills MSMU graduates obtain are also obtained by the certificate participants.

With the prevalence of data across all disciplines, the demand for data literate professionals extends beyond the traditional computational STEM career paths. The goal of this certificate program is to provide a foundation in STEM concepts to certificate program participants regardless of prior technology background. As the target participants will be from diverse discipline backgrounds, the program is designed to leverage comparable advisory relationships beyond the School of Natural Science and Mathematics to include the College of Liberal Arts, and the Bolte School of Business. Utilization of a cross-disciplinary faculty panel specific to the capstone project further enables the program to consider a broad spectrum of career directions for graduates from the program.

The data science applications course within the program is designed to provide participants with examples of data science applications directly tied to the technology skills learned in foundational courses. Data model applications covered will extend beyond applications related to STEM career paths only, and include a diversity of application areas taught by faculty with application experiences across all disciplines.

# D. Quantifiable Evidence of Market Supply & Demand

In a frequently referenced article in 2011, McKinsey Global Institute (MGI) and McKinsey's Business Technology Office indicated "...The amount of data in our world has been exploding, and analyzing large data sets—so called big data—will become a key basis of competition, underpinning new waves of productivity, growth, innovation, and consumer surplus..."

Additionally, the report indicated "...There will be a shortage of talent necessary for organizations to take advantage of big data. 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 managers and analysts with the know-how to use the analysis of big data to make effective decisions..."

These same observations were restated in the March, 2017 <u>Inside Higher Ed</u> article in which the shortage of data literate job candidates was highlighted as among the nation's top skills gaps.

In 2016, as Mount St. Mary's was evaluating direction specific to Data Science, a market research study was conducted by Illumine8 Marketing and PR. Results of that study provided a summary of job demands within regions of the country specific to data science. Highlighted on

the report were the Washington-Northern Virginia area consistently within the top 5 and most recently at #3, along with the Baltimore-Towson area at #11.

Alternate Location 3	Job Volume :	Salary Range	Candidate Supply	Posting Period	Hiring Scale
San C. Sandyvalo Santa Clara, CA	2,716	E37K - 154K	18,000	49 da, s	E 1775
San i Concisco-Del "Sad- Pracco di GA	2,370	132K-160K	21.006	47. da; s	E 3137
Vissbirgton-Adington- Alexardna, ÖÖ	2,253	\$82K - 144K	15,000	49 đạ, ş	# <b>  1</b> 78
nas a York-dantin in New Jersey-Long Island, 174	1,565	\$20K-148K	27,000	44 days	<b>1</b> 78
Douter-Combinde- Grack, N. A.	1,181,	,573K-138K	13,000	50 days	£ <b>]</b> 76
ensitie-than na-Hallayde, WA	1,089	978K+126K	14,006	ភូមិ ជន៖ រ	# <b>#</b> 81
Bully Fort Joya Amily To, XX	776	555K+128K	3,200	42 dayr	<u>1</u> 1177
Chicago-JoSayCapena 🖔 . ".	721	372K-137K	31,000	40 da, s	D: 11.79
Los Angeles-Long Brack- Ganta Anal CA	641.	\$50K-120K	12.006	47 d <b>o</b> ja	F. 10
EpingheCter, LC	519	\$84K-125K	2,500	47 ds;:	ji: <b>18</b> 78

Top Geographic Data Science Job Markets

Frederick County is the third fastest growing county in Maryland with a .9% increase in population from July, 2015 to July, 2016. With its location equidistant to the Washington as well as the Baltimore metropolitan areas, the county has many professionals working in those regions across a broad spectrum of careers. Government is a big presence in both metropolitan areas, and a large employer of residents within Frederick County. Mount St. Mary's University is well positioned with a Frederick and Emmitsburg campus to reach the non-traditional adult students within the Frederick County communities. Additionally, government locations like FEMA in Emmitsburg, and Ft. Detrick in Frederick are local employers with heavy reliance on data literate professionals.

Using the Indeed job engine to conduct search within the Maryland job market for Data Scientist, produced a listing of 954 current positions on May 23, 2017. This included positions in Frederick, Rockville, and Columbia among many other regional locations. Using the Indeed job engine to conduct comparable search within the Maryland job market for Data, produced a listing of 14,603 positions.

#### **Educational and Training Needs**

Focus in most educational institutions has been on providing the graduate-level training for the data scientist roles. Data literate employees are needed across a spectrum of disciplines, and job titles. Many of these positions do not require the graduate level understanding of statistics,

computer science, and programming skills that Data Science programs address. The needs are more in terms of comfort with the variety of formats of data, how to acquire data, how to clean and engineer data, how to prepare data for decision making, and how to appropriately leverage that data in the decision-making process.

The certificate program proposed is designed to address that gap. Program participants would receive a level of training in the technology skills that would prepare them as a valuable employment asset within their current discipline. As explained in more detail in section E below, there are no comparable certificate programs within Maryland. The programs that do exist have significant prerequisite programming, computer science, and/or statistics expectations for candidates to be eligible to apply into the program.

According to a combination of U.S. Census Bureau data as well as Maryland Report Card data, Frederick County has a 92% high school graduation rate; second in the state of Maryland behind Howard County at 95%. Thirty-nine (39) % of residents have a bachelor's degree or higher; third in the state of Maryland. Of Frederick County high school graduates, 76.7% enroll in college within 24 months of high school graduation; 71% of the population age 16 or older are in the civilian work force. Data support the focus that career education and advancement represent to Frederick County residents. Web sources referenced include 2016 Maryland Report Card and TownCharts.

Estimating potential student populations is difficult as Data Science programs are not highly common at this time, and enrollment numbers not easily obtained as a basis for comparison. Census data indicates that by the year 2020, Frederick County will have 67,351 individuals in the non-traditional college age bracket (age 25-44). Applying the percent of residents with college degree against this population, we obtain 26,267 individuals in Frederick County with educational background to participate in certificate program. The program aims to enroll 30 students in initial courses representing .1% of the total number of eligible individuals.

We also anticipate internal demand for this certificate program related to current graduate students across the MSMU curriculum; a population of 341 graduate students is currently enrolled.

# E. Reasonableness of Program Duplication

A search of the <u>Maryland Higher Education Commission site</u> indicates there are no bachelor's degree programs in Data Science, nor any post-baccalaureate certificate programs. Proposals submitted within the last 2 years were found for Loyola University related to a master's degree in Data Science. Comparable Data Science proposal was found for Johns Hopkins. Johns Hopkins University offers a post-masters certificate (PMC) in data science as does the University of Maryland.

Johns Hopkins admissions requirements include traditional computer science baccalaureate courses in programming languages, discrete mathematics, data structures, and calculus. Loyola

has admissions requirements that include statistics course and indication of mathematics maturity. UMD PMC admissions requirements seek programing language proficiency demonstrated through coursework or software development experience. Several computer science departments offer data science as a minor or specialized track within the bachelor's degree curriculum.

Specific to the western Maryland area with Frederick County, there is no duplication of program. Specific to the other data science programs highlighted above as well as many online curriculum, the focus is on individuals who have completed some level of prior training in computer science and computational mathematics. No programs with focus on attracting individuals to data science without the constraint of prior programming knowledge were identified.

The table below compares the proposed MSMU PBC curriculum to existing data science certificate programs within the state of Maryland. This table highlights the differences related to target audience and pre-requisite requirements as well as overall requirements to complete certificate.

Anst.	Gradiate Boream Science – Graduate Level	PostaBaccatanreau S CAP attente Coption & Proposed	Bachelor's degree in any discipline, no programming or statistics pre-reqs	1) Introduction to Data Science 2) Data Engineering 3) Predictive Model Algorithms 4) Data Science applications 5) Data Science capstone project
Loyola University of Maryland	M.S. Data Science	No	Introductory statistics, mathematical maturity judged by admissions committee, introduction to computer science/programming	N/A
Johns Hopkins	PMC Data Science M.S. Data Science	Yes, PMC not PBC	Multivariate calculus; discrete mathematics or linear algebra or differential equations; courses in Java or C++ or Python; course in data structures	6 courses; at least 2 must be from Applied and Computational Mathematics area and 2 from Computer Science area; at least one must be 700 level
UM College Park	Certificate Data Science	Yes, PBC	Prior coursework establishing quantitative ability (i.e. calculus, linear algebra, basic statistics etc.). Proficiency in programming languages, demonstrated either	4 courses; Machine Learning, Statistics, Databases, Visualization

			through prior programming coursework or substantial software development experience.	
Towson University	M.S. Computer Science	No	N/A	N/A
UMBC	M.S. Computer Science	No	N/A	N/A
Hood	M.S. Computer Science	No	N/A	N/A

# F. Relevance to Historically Black Institutions (HBIs)

### 1. Potential impact on high-demand programs at HBI's

Not applicable. The proposed program does not duplicate or compete with programs at any of the regional HBIs.

#### 2. Potential impact on the uniqueness and missions of HBIs

Not applicable. The proposed program does not duplicate or compete with programs at any of the regional HBIs.

### G. Evidence of the Principles of Good Practice for Online Programs

Not applicable.

# H. Adequacy of Faculty Resources

The Data Science certificate program will be housed administratively within the Mathematics and Computer Science Department. The faculty who will teach foundational courses in the program will come from that department. These faculty have demonstrated their ability to teach computer science and data science concepts within the curriculum. A computer scientist, with graduate degree and some data science experience, will be hired to provide additional knowledge and expertise related to this program going forward. In the first iterations of the Data Science certificate program, faculty teaching the applications course and participating on the capstone panel will come from a small cross-disciplinary group highlighted below. As the program matures, the expectation is that many other application areas aligned to the foundational curriculum will be developed and integrated into the applications course.

The majority of the faculty listed below have terminal degrees in their field and all are full-time members of the university faculty.

# Data Science Foundational Computer Science Courses

Rebecca Portier, M.S. in Computational Mathematics – Ms. Portier has extensive experience in industry applications related to data and data models including roles as Operations Research Analyst, Mathematician, and Technology Manager. Prof. Portier teaches Computer Science and Data Science curricula. [Assistant Professor Computer Science, Full-time]

Brian Heinold, Ph.D. in Mathematics – Dr. Heinold has taught across the Mathematics, Computer Science and Cybersecurity curricula [Associate Professor Mathematics and Computer Science, Full-time]

Frederick J. Portier, Ph.D. in Mathematics—Dr. Portier has taught across the Mathematics, Computer Science and Cybersecurity curricula including early iterations of the current Data Science minor courses [Professor Mathematics and Computer Science; Full-time]

Scott Weiss, M.S. in Computer Science–Mr. Weiss has taught across the Computer Science curricula including Artificial Intelligence and Theory of Computation [Assistant Professor Computer Science, Full-time]

#### **Applications Course**

Dr. Caitlin Faas, Ph.D. in Human Development – Dr. Faas has taught across the Psychology curricula including Research Preparation and Experimental Cognition [Assistant Professor of Psychology, Full-time]

Dr. Abigail Kula, Ph.D. in Behavior, Ecology, Evolution and Systematics—Dr. Kula has taught across the Environmental Science and Biology curricula including Biostatistics and Ecology [Assistant Professor Environmental Science, Full-time]

Dr. Layton Field, Ph.D. in Sociology – Dr. Field specializes in population research and has taught across the Sociology/Criminal Justice curricula including Population and Society, Social Research Methods, Statistics, Introduction to Criminal Justice, and Social Problems [Assistant Professor Sociology and Criminal Justice, Full-time]

Over time, this list will expand to include application courses by full-time faculty members across the curricula.

# I. Adequacy of Library Resources

Mount St. Mary's University's Hugh J. Phillips Library currently contains about 200,000 bound volumes and a rapidly expanding collection of scholarly information databases that provide convenient access to e-books, journal articles and a variety of data sources. Included in our e-

library are more than 25,000 professional and scholarly journal publications that are carefully chosen to support each of the University's academic programs.

The library has an excellent E-resources collection that includes discipline specific databases including the complete JSTOR back files. Content from Sage, EBSCO, ProQuest, Duke e-journals, ATLA and many others is available from the library's website <a href="http://libguides.msmary.edu/phillipslibrary">http://libguides.msmary.edu/phillipslibrary</a>. The library recently implemented the *EBSCO Discovery Service* that performs a single search of all library resources from one search interface. Computer Science, Applied Mathematics, Statistics, and many discipline specific subscriptions are currently available. Requests for additional resources can be made each year.

Our library staff includes four faculty librarians who provide research assistance and information literacy instruction to individuals and groups. A faculty librarian with theological training maintains the theology collection of approximately 46,000 volumes. Our main desk services, resource acquisitions, cataloging and interlibrary loans are provided by four student/faculty-focused employees, with the help of several dedicated student assistants.

The Phillips Library is a founding member of the Maryland Interlibrary Consortium and collaborates with Hood College, Baltimore International College, Washington Adventist University (formerly Columbia Union College), Loyola College-Notre Dame University Library, and Stevenson University. Through this consortium, Mount students and faculty have direct access to the collections of each member library through electronic and physical delivery services. The average delivery time for print materials is within 24hours.

Table 1. 2015-16 Library Expenditures	
Volumes	149,287
Per FTE student	72
Journal Titles-Paper	233
Journal Titles-Digital	26,544
Librarian Research Transactions	1.068
Participation in Instruction Services	1,210
Databases	130
Videos	1,500
Total Library Expenditures	\$862,061
Library expenditures per FTE student	\$ 413

Source: Mount St. Mary's Factbook 2017

# J. Adequacy of Facilities, Infrastructure and Instructional Equipment

The Data Science program will be offered through MSMU's Emmitsburg and Frederick Campus locations. As part of the Mathematics and Computer Science department, the Data Science

certificate program will have access to the Coad Science Building facilities utilized by the program. Coad is a 48,000 ft<sup>2</sup> building that holds classrooms, faculty and staff offices, specialized laboratories, a vivarium, a computer lab, and a greenhouse. Existing computer laboratory space was recently renovated to expand utilization for the new Cybersecurity major. Those renovations were done with data science needs in mind as well, and will provide access to servers that can be leveraged for any custom software or data access needs by the program.

The Frederick Campus is a 25,000 ft<sup>2</sup> facility with classrooms, offices, large conference room, two dining areas, chapel, and kitchen. The facility has some unused capacity in terms of classroom space so it will support the 1-2 additional courses per term that the Data Science program will introduce. Although it is a technical program, laboratory facilities are not needed at this site. Faculty instructors have access to the full resources of the facility including photocopiers, scanners, audio-visual equipment, phones, and office supplies. Administrative assistants provide administrative support and faculty also may avail themselves of the resources of the MSMU Career Center, Learning Services, Information Technology Support Center, and Health and Wellness Center.

Numerous online data and programming resources are available and will be leveraged throughout the curriculum. Participants will be expected to provide their own computer. Software leveraged will be predominately open source programming languages such as Python and R with integration to a variety of readily available open source tools like Spark, Hadoop, Shiny, MySQL, etc. Governments across the country and around the world have numerous OpenData database sources readily available for access and analysis. Interfaces with existing websites through API's, including a variety of social media sites, will be leveraged by participants.

In summary, this program can be offered with existing institutional facilities, infrastructure, and instructional equipment.

# K. Adequacy of Financial Resources

TABLE 2: RESOURCES						
Resources Categories	Year 1 (2017-2018)	Year 2 (2018-2019)	Year 3 (2019-2020)	Year 4 (2020-2021)	Year 5 (2021-2022)	
1. Reallocated Funds	\$0	\$0	\$0	\$0	\$0	
2. Tuition/Fee Revenue (c+g)	\$0	\$117,720	\$161,760	\$249,840	\$321,750	
a. # F.T. Students						
b. Annual Tuition/ Fee Rate (Discounted rate)						
c. Annual Full Time Revenue					· · · ·	
(a x b)						
d. # Part Time Students	0	12	16	24	30	
e. Credit Hour Rate	\$635	\$654	\$674	\$694	\$715	
f. Annual Credit Hours	0	15	15	15	15	

g. Total Part Time Revenue (d	\$0	\$117,720	\$161,760	\$249,840	\$321,750
x e x f)					
3. Grants, Contracts, & Other					
External Sources					
4. Other Sources					
TOTAL (Add 1-4)	\$0	\$117,720	\$161,760	\$249,840	\$321,750

Academic year 2017-2018 will be primarily a year of preparation because we will not have time to recruit students. After a strong recruiting and marketing effort in 2017-18, we expect to admit a first cohort of 12 students in the fall of 2018. We expect to grow by about 30% after the first cohort, reaching 50% increase in subsequent years, and leveling off at about 30 students.

<u>Credit Hour Rate</u>: The rate for 2016-17 is \$635 per credit. We project an increment of 3% per year which is a typical amount of increase at MSMU.

<u>Total Resources</u>: The resources available are projected to be \$0 in year 1, increasing to \$321,750 in year 5.

TABLE 3: EXPENDITURES					
Expenditure Categories	Year 1	Year 2	Year 3	Year 4	Year 5
1. Faculty (b+c below)	\$0	\$113,900	\$116,178	\$118,502	\$121,485
a. # FTE	0	1	1	1	1
b. Total Salary	\$0	\$85,000	\$86,700	\$88,434	\$90,203
c. Total Benefits	\$0	\$28,900	\$29,478	\$30,068	\$31,282
2. Admin. Staff (b+c below)					
a. # FTE					
b. Total Salary					
c. Total Benefits					
3. Support Staff (b+c below)	\$3,350	\$3,417	\$3,485	\$3,555	\$3,626
a. # FTE	0.05	0.05	0.05	0.05	0.05
b. Total Salary	\$2,500	\$2,550	\$2,601	\$2,653	\$2,706
c. Total Benefits	\$850	\$867	\$884	\$902	\$920
4. Equipment					
5. Library		\$2000	\$2100	\$2200	\$2300
6. New or Renovated Space					
7. Other Expenses (see Table 3)	\$10,200	\$9,700	\$9,700	\$9,700	\$9,700
8. TOTAL (Add 1 – 7)	\$13,550	\$129,017	\$131,463	\$133,957	\$137,111

<u>Faculty</u>: By AY2018-2019, we will hire a full-time, tenure-track faculty member in computer science to assist in teaching the data science courses along with courses in support of existing computer science programs. This will result in a full teaching load (21 credits per year). The range of median salaries mid-level data science professionals is \$74,000 - \$140,000 (Payscale; Bureau of Labor Statistics). The national average for Asst. Professor Computer Science faculty (first-year) is \$84,281 (Inside Higher Ed). In order to be able to attract and retain an experienced individual in this field a

salary of \$85,000 is recommended. A portion of time has been allotted in AY2017-2018 for current faculty to begin development of curriculum in preparation for course offerings in fall of 2018.

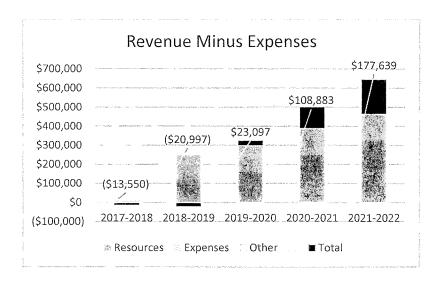
<u>Support Staff</u>: We estimate a time commitment equivalent to 5% of a person's workload in the Communications Office for marketing and promotion. A salary of \$50,000 was assumed and benefits are 34% of the salary. The salary was incremented by 2% per year.

<u>Library</u>: To purchase some additional journal subscriptions, \$2,000/year should be added to the Library budget. This amount was incremented by 5%/year to account for inflation.

Other Expenses: See table 3 below.

TABLE 3: OTHER EXPENSES						
Promotional items	\$1,000	\$500	\$500	\$500	\$500	
Printed brochures/flyers	\$200	\$200	\$200	\$200	\$200	
Advertising/Marketing expenses	\$5,000	\$3,000	\$500	\$300	\$300	
Faculty development		\$2,000	\$2,000	\$2,000	\$2,000	
Memberships	\$1,500	\$1,500	\$1,500	\$1,500	\$1,500	
Conferences/travel	\$1,500	\$1,500	\$1,500	\$1,500	\$1,500	
Networking/outreach	\$1,000	\$1,000	\$1,000	\$1,000	\$1,000	
Subtotal	\$10,200	\$9,700	\$7,200	\$7,000	\$7,000	

The "Other Expenses" include operating expenses for the program such as promotional items (pens, lanyards, etc.), printed brochures, and marketing and advertising (such as online advertising). Funds will be available for faculty development (e.g., travel to conferences) so that they will remain up-to-date on current practices. Department will obtain membership in relevant societies and will have funds for traveling to conferences. Finally, funds for networking and outreach will enhance recruitment.



Budget Summary: In the first 2 years of the program, the program runs at a loss while marketing and building full cohort. In year 3, there is a small amount of profit, while revenue far exceeds expense in all subsequent years.

# L. Adequacy of Provisions for Evaluation of Program

The program will leverage significant curriculum from the existing Data Science minor along with data science applications courses across discipline in order to tailor curriculum for the graduate certificate. The program will be part of the Middle States Accreditation of the university. Course evaluations will be completed for each course as designated by the College/School in which the course resides and the university. Full-time faculty are reviewed at least every five years. Part-time faculty are reviewed on a course/semester basis. Each program is reviewed every five years, using an outside consultant. The following table details department Learning Outcomes to be assessed at least once in a five-year period.

Learning Outcome	Assessment	Benchmark	Timing
LO1: demonstrate an understanding of the lifecycle processes related to data science	Rubric assessment through cross- disciplinary panel within capstone	TBD	Every spring in Capstone
LO2: have the ability to apply the tools and techniques of data engineering and statistics to effectively investigate and solve real-world problems	Rubric assessment through cross- disciplinary panel within capstone	TBD	Every spring in Capstone
LO3: have the ability to communicate technical ideas with precision and clarity; demonstrate ability to select appropriate data models/applications related to the problem domain.	Rubric assessment through cross- disciplinary panel within capstone	TBD	Every spring in Capstone
LO4: understand the ethical issues	One and five-year surveys of alumni	TBD	Every spring

that constitute the		
data science		
profession so that		
they are prepared		
for success in a		
career or graduate		
study		

# M. Consistency with the State's minority student achievement goals

The Data Science certificate program at MSMU will be promoted along with other graduate programs in SNSM. In 2015-16, the proportion of students of color was 20% in the graduate programs and 30% in the undergraduate programs. Our commitment to diversity is evidenced by a recent S-STEM award from the National Science Foundation that provides scholarship funding for underrepresented students in STEM majors with high financial need.

#### **Nondiscrimination Statement**

It is the policy of Mount St. Mary's University not to discriminate on the basis of race, color, national or ethnic origin, political or religious opinion or affiliation, age, sex or handicapping condition in the recruitment or admissions of students, or in the administration of the university's educational policies, admissions policies, scholarship and athletic programs, and other university-administered activities and programs.

### **Center for Student Diversity**

The Center for Student Diversity was established to aid Mount St. Mary's University in its efforts of fostering inclusion, collaboration, and relationship building across campus. The Center provides academic, social, and transitional support in addition to programming, leadership training and inclusive workshops for ALL students and promotes exchange and dialogue between individuals of diverse backgrounds.

The Center for Student Diversity oversees the intercultural development programs, the Horning Fellowship, student support programs (including Third Century Scholars program and the American Indian program), and cultural programs. The office also supports cultural organizations, conducts diversity awareness programs, assesses the needs and climate of diverse groups and advocates on behalf of underrepresented students.

# N. Relationship to low productivity programs identified by the Commission

Not applicable. There are no identified low productivity programs at MSMU.



# **Data Science Tables**

# K. Adequacy of Financial Resources

TABLE 2: RESOURCES									
Resources Categories	Year 1 (2017-2018)	Year 2 (2018-2019)	Year 3 (2019-2020)	Year 4 (2020-2021)	Year 5 (2021-2022)				
1. Reallocated Funds	\$47,720	\$24,602	\$0	\$0	\$0				
2. Tuition/Fee Revenue (c+g)	\$0	\$117,720	\$161,760	\$208,200	\$321,750				
a. # F.T. Students	\$0	\$0	\$0	\$0	\$0				
b. Annual Tuition/ Fee Rate (Discounted rate)	\$0	\$0	\$0	\$0	\$0				
c. Annual Full Time Revenue (a x b)	\$0	\$0	\$0	\$0	\$0				
d. # Part Time Students	0	12	16	24	30				
e. Credit Hour Rate	\$635	\$654	\$674	\$694	\$715				
f. Annual Credit Hours	0	15	15	15	15				
g. Total Part Time Revenue (d x e x f)	\$0	\$117,720	\$161,760	\$249,840	\$321,750				
3. Grants, Contracts, & Other External Sources	\$0	\$0	\$0	\$0	\$0				
4. Other Sources	\$0	\$0	\$0	\$0	\$0				
TOTAL (Add 1-4)	\$47,720	\$142,322	\$161,760	\$208,200	\$321,750				

TABLE 3: EXPENDITURES								
Expenditure Categories	Year 1	Year 2	Year 3	Year 4	Year 5			
1. Faculty (b+c below)	\$34,170	\$113,900	\$116,178	\$118,502	\$121,485			
a. # FTE	.3	1	1	1	1			
b. Total Salary	\$25,500	\$85,000	\$86,700	\$88,434	\$90,203			
c. Total Benefits	\$8,670	\$28,900	\$29,478	\$30,068	\$31,282			
2. Admin. Staff (b+c below)	\$0	\$0	\$0	\$0	\$0			
a. # FTE	\$0	\$0	\$0	\$0	\$0			
b. Total Salary	\$0	\$0	\$0	\$0	\$0			
c. Total Benefits	\$0	\$0	\$0	\$0	\$0			
3. Support Staff (b+c below)	\$3,350	\$3,417	\$3,485	\$3,555	\$3,626			
a. # FTE	0.05	0.05	0.05	0.05	0.05			
b. Total Salary	\$2,500	\$2,550	\$2,601	\$2,653	\$2,706			
c. Total Benefits	\$850	\$867	\$884	\$902	\$920			
4. Equipment	\$0	\$0	\$0	\$0	\$0			
5. Library	\$0	\$2000	\$2100	\$2200	\$2300			
6. New or Renovated Space	\$0	\$0	\$0	\$0	\$0			
7. Other Expenses (see Table 3)	\$10,200	\$9,700	\$9,700	\$9,700	\$9,700			
8. TOTAL (Add 1 – 7)	\$47,720	\$129,017	\$131,463	\$133,957	\$137,111			