University of Kansas Medical Center

MS in Health Data Science

Program Approval

I. General Information

A. Institution University of Kansas Medical Center

B. Program Identification

Degree Level:	Masters
Program Title:	Health Data Science
Degree to be Offered:	Master's in Health Data Science
Responsible Department or Unit:	University of Kansas/School of Medicine/Department of Biostatistics & Data Science
CIP Code:	30.7001
Modality:	Hybrid
Proposed Implementation Date:	Fall 2022

Total Number of Semester Credit Hours for the Degree: 36

II. Clinical Sites: Does this program require the use of Clinical Sites? No

III. Justification

The Master's Degree in Health Data Science is proposed by the Department of Biostatistics & Data Science at the University of Kansas Medical Center with full support of the leadership of the University of Kansas Medical Center. This online and in-person program will be on the cutting edge for several reasons. It is the first in the region focused on producing graduate-trained health data scientists with the high level of biostatistics and computing skills demanded by a rapidly emerging healthcare analytics workforce. In addition, online access to all required coursework provides flexibility to accommodate working professionals seeking advanced health data science training.

Due to the advent of new technologies, a large amount of data is being generated in healthcare industries. For example, new sensor technologies have dramatically increased the frequency and reliability of the data being generated by individual patients. Therefore, there is a high demand for expertise in tracking, managing, analyzing, and interpreting the high volume of data being generated. In addition, interdisciplinary research is a point of emphasis in academia, government, and industry. To support effective interdisciplinary collaborations, data scientists need to possess statistical, computing, and domain-level expertise. The application of data science has already had a tremendous impact on the diagnosis and treatment of many medical conditions. However, clinicians/researchers usually do not have formal training in data science in their degree programs. The proposed MS in Health Data Science will fill this significant gap.

Besides clinicians and researchers already working in their field, the proposed program will be an attractive career choice for beginners. Integrating clinical data with other diverse data sources, like sleep cycles, sedentary vs active life, diet, and nutrition, etc., allows a greater level of understanding of the association and casual factors. Moreover, each of us carries diverse genetic variants and lifestyle factors that can be combined with phenotypic and demographic data to inform our understanding of physiology, which will eventually help in advancing personalized medicine. Analyzing the vast amount of available data to generate actionable information for clinicians, requires advanced training in health data science.

Highly personalized data comes with a unique management challenge due to the Health Insurance Portability and Accountability Act (HIPPA). Both law and ethics require that the patients' personal information must be kept secure. On the other hand, utilization of as much clinical research data as possible is vital for the development of personalized medicine. The complexity of securing such information and using it effectively demands highly skilled health data scientists immersed in the healthcare world.

The curriculum has been developed by the largest group of PhD statisticians and data scientists in the KU system, whose track record of successful online curriculum development and instruction is evidenced by the rapid growth of their existing programs. The proposed degree will deliver high-quality accessible graduate programs in STEM fields and produce a much-needed health data science workforce for the local and regional communities of Kansas City and the State of Kansas.

IV. Program Demand:

A. Survey of Student Interest

Number of surveys administered: ……	N/A
Number of completed surveys returned: …….	137
Percentage of students interested in program: …	39.4%

An electronic RedCAP survey was distributed through the KUMC Office of Graduate Medical Education, the Office of Postdoctoral Affairs and Graduate Studies, the University of Kansas Cancer Center, and Frontiers. In addition, the survey was distributed to select regional academic institutions. The survey first asked participants to indicate whether a Master of Science in Health Data Science degree was a career goal, and for those who answered in the positive, a series of questions about the types of degrees desired were presented. Of the 137 respondents, 29 (21.2%) were undergraduate students, two (1.5%) were medical students, 50 (36.5%) were graduate students, 30 (21.9%) were medical residents, nine (6.6%) were medical doctor, 16 (11.7%) were post-doctoral fellows, three (2.2%) were staff, and five (3.6%) were faculty. In total, 54 (39.4%) indicated health data science was a career goal. All 54 (100%) replied "Yes" to the question "Would you be interested in a Master of Science degree in Health Data Science at the University of Kansas Medical Center?"

B. Market Analysis

The Master of Health Data Science is designed to prepare students for immediate entry into the workforce. "Data Scientist" has been ranked as one of the top jobs in the U.S. over the last several years. The Harvard Business Review named it as "the sexiest job of the 21st century" back in 2012 and pointed out that the shortage of data scientists is becoming a serious impediment in some sectors (Davenport & Patil, 2012). The recruiting website, Glassdoor, annually releases the best jobs in U.S. based on three criteria: earning potential (median annual base salary), overall job satisfaction, and the number of job openings on Glassdoor. Glassdoor ranks "Data Scientist" as the second best job in U.S. in 2021 with median salary of $113,736. Further, "Data Scientist" has been consistently ranked as top three best job in last five years (third in 2020, first in 2019, 2018, 2017, and 2016) (Glassdoor, 2020). Bureau of Labor Statistics determined that data science is one of the fastest growing occupations and has a projected 31% growth over the next 10 years (2019-2029) (Bureau of Labor Statistics, 2020). However, the supply of data scientists has not been able to keep pace with their demand. In August 2018, LinkedIn reported that there was a shortage of 151,717 people with data science skills in the United States (LinkedIn, 2018). A January 2019 report from Indeed showed a 29% increase in demand for data scientists every year and a 344% increase since 2013. The same report also showed that searches for data science jobs increased 14% in 2018, which suggests a big gap between the demand and supply (Flowers, 2019).

In the Kansas City Metro area, the demand for workers with statistics and analytics skills will increase by 22.2% over the next decade, according to a Labor Market report by the Regional Workforce Intelligence Network of

Greater Kansas City (MARC, 2020). According to the Bureau of Labor, the projected demand for Statisticians and Data Scientists will be 34.6% and 31% respectively. The proposed program will be a combination of both statistics and data science with a focus on health outcome data. Therefore, the program will be in high demand.

There are several factors that influence the trend in healthcare analytics market size and growth. The transition from paper charts to real-time monitoring systems and use of electronic health records to gather patient health data is expected to increase the healthcare analytics market size. Industry players invest huge amounts of money in research and development processes to create unique platforms and solutions with enhanced features that allow them to gain a competitive advantage in the market for health care analytics. This increase in investment is expected to fuel the growth of the healthcare analytics market size. In addition, the outbreak of COVID-19 pandemic has brought the importance of data analytics sharply into focus. The need to extract and analyze the healthcare data quickly has only increased with pandemic (Sheng, 2020). Integrating the developments in computing technology and increasing the implementation of predictive and prescriptive analytics in most hospitals is driving healthcare analytics market growth.

V. Projected Enrollment for the Initial Three Years of the Program

Based on the demand and the current recruitment in other MS in Applied Statistics programs, we anticipate the following estimates of the enrollment in the proposed program.

Year	Headcount Per Year		Sem Credit Hrs Per Year	
	Full- Time	Part- Time	Full- Time	Part- Time
Implementation	5	5	95	45
Year 2	10	10	180	90
Year 3	10	10	180	90

VI. Employment

Implementation of the program will create several employment opportunities for both administrative staff and faculty, as well as graduates. There will be a Director, an Assistant Director, and an Administrative Assistant for the program. The directors will have dedicated percentage effort time overseeing and ensuring the proper functioning of the program. The full-time Administrative Assistant will work on administrative processes and facilitate instructor and trainee communication, starting from student recruitment to graduation. Many other KUMC faculty will be employed with percentage effort for their involvement in various educational activities including both didactic and non-didactic course works.

There will be a wide spectrum of employment opportunities for the graduates of the proposed program. Given the applied nature of coursework/training, we anticipate that our graduates will be able to join the workforce immediately after the graduation. The program will bring statistics and data science together with a focus on health data. This combination of skill sets is highly sought after and required in many healthcare institutions and industries. We anticipate that our graduates will be targeted for recruitment by academic health research institutions and a variety of different health care provider industries. A few examples of potential employment venues include academia, government, hospital, biotechnology, insurance company, pharmaceutical company, and consulting. Many of the healthcare analytics providers are increasingly focused on providing scalable solutions that can be made adaptable to technology advancement as well as human expertise and skills in healthcare organizations. Examples of a few companies include, Cerner Corporation, PRA Healthcare, Quintiles, McKesson Corporation, Optum, CitiusTech, Health Catalyst, SAS Institute Inc, and VitreosHealth, Inc. Therefore, successful completion of the proposed MS in Health Data Science program will enable our graduates achieve high level of competency to seek a career across a wide spectrum of healthcare data scientist positions.

VII. Admission and Curriculum

Admission to the MS program in Health Data Science is made by the Department's MS/PhD Admissions Committee which consists of four full-time faculty members and the Program Director who serves as Chair. All committee recommendations are presented directly to the Department Chair. ***Non-discrimination policy*** : Qualified students are admitted without regard to race, color, ethnicity, religion, sex, national origin, age, ancestry, disability, status as a veteran, sexual orientation, marital status, parental status, gender identity, gender expression, and genetic information.

A. Admission Criteria

Application for admission to the MS program in Health Data Science is through the Department of Biostatistics and Data Science at the University of Kansas Medical Center. The following materials are required to be considered for admission:

- A completed online application form.
- Evidence of a bachelor's degree from an accredited college or university.
- Official transcript(s) bearing the official seal from each college or university in which course work had been taken.
- Three (3) letters of recommendations.
- TOEFL or IELTS test results (if applicable).

For admission into the MS program in Health Data Science, the applicant must meet the general requirements for admission to Graduate Studies. The minimum requirements for admission into the MS program are:

- A cumulative undergraduate GPA of 3.0 or better is required for regular admission status. An applicant with an undergraduate GPA not meeting the minimum requirements may be admitted under provisional status, provided they meet the academic standards of the Department without excessive deficiencies in pre-requisites.
- B average (or higher) in Calculus I – II (i.e., single variable differentiation and integration or equivalent).
- Successful completion of a course in any computer programming language.

An applicant meeting the minimum requirements for admission is referred to the Admissions Committee for approval or disapproval. Approval for admission is good for up to 12 months from the approved date for admission. Failure to enroll during this time will require a new application if admission is desired.

B. Curriculum

The proposed educational program will utilize all but one existing statistical, computational and health data science foundation courses. One new course will be designed with special focus on observational health data analysis. The curriculum of the MS in Health Data Science is built upon three foundational courses: required statistics foundation course (12 credit hours), required computing foundation courses (six credit hours), required health data science foundation courses (12 credit hours). In addition, the program requires six credit hours of elective courses of students' choice.

Required Statistics Foundation Courses (12 semester credit hours (SCH))
HDSC 805: Professionalism, Ethics and Leadership in the Statistical Sciences (3 SCH)
HDSC 835: Categorical Data Analysis (3 SCH)
HDSC 840: Linear Regression (3 SCH)
HDSC 845: Survival Analysis (3 SCH)

Required Computing Foundation Courses (6 semester credit hours (SCH))

HDSC 818: Introduction to R (1 SCH)
HDSC 819: Introduction to Python (1 SCH)
HDSC 822: Introduction to SQL (1 SCH)
HDSC 823: Introduction to Programming and Applied Statistics in R (3 SCH)

Required Health Data Science Foundation Courses (12 semester credit hours (SCH))

HDSC 824: Data Visualization and Acquisition (3 SCH)
HDSC 880: Data Mining and Analytics (3 SCH)
HDSC 881: Statistical Learning I (3 SCH)
HDSC 861: Observational Health Data Analysis (3 SCH)

Electives (6 semester credit hours (SCH))
HDSC 815: Introduction to Bioinformatics (3 SCH)
HDSC 820: SAS Programming I (3 SCH)
HDSC 830: Experimental Design (3 SCH)
HDSC 855: Statistical Methods in Genomic Research (3 SCH)
HDSC 882: Statistical Learning II (3 SCH)

Course Plan

The courses mentioned above will be taught in Fall, Spring and Summer semesters as shown on the plan of study below:

Year 1: Fall **SCH = Semester Credit Hours**

Course #	Course Name	SCH
HDSC 835	Categorical Data Analysis	3
HDSC 840	Linear Regression	3
HDSC 824	Data Visualization and Acquisition	3

Year 1: Spring

Course #	Course Name	SCH
HDSC 880	Data Mining and Analytics	3
HDSC 845	Survival Analysis	3
HDSC 818	Introduction to R	1

Year 1: Summer

Course #	Course Name	SCH
HDSC 823	Introduction to Programming and Applied Statistics in R	3

Year 2: Fall

Course #	Course Name	SCH….
HDSC 881	Statistical Learning, I	3
HDSC 822	Introduction to SQL	1
HDSC 815 (1st Elective)	Introduction to Bioinformatics (or one other elective)	3

Year 2: Spring

Course #	Course Name	SCH….
HDSC 861	Observational Health Data Analysis	3
HDSC 882 (2nd Elective)	Statistical Learning II (or one other elective)	3

Year 2: Summer

Course #	Course Name	SCH….
HDSC 805	Professionalism, Ethics and Leadership in the Statistical Sciences	3
HDSC 819	Introduction to Python	1

Total Number of Semester Credit Hours ……. 36

All the courses required for the MS in Health Data Science already exist as part of the MS in Applied Statistics and MS in Biostatistics programs with one exception. One new course (HDSC 861) will be created to align with the Health Data Science degree focus. The courses will be taught both in-person and online. The online courses will be delivered via internet and require the students to have access to an adequately equipped computer with internet connection. The Department of Biostatistics & Data Science currently offers online courses in two formats: synchronous (live) or asynchronous (recorded). In general, the Department of Biostatistics & Data Science has developed online courses to mimic the classroom as closely as possible using educational technology, such as, Panopto, Zoom, Blackboard, and Canvas.

Annual Evaluations

Students will be evaluated each May by their faculty advisor and Program Chair. These evaluations provide feedback to the student regarding the progress that they are making towards the degree. To proceed in the program, students must remain in good standing and maintain a 3.0 GPA with no more than two grades of C in the required courses.

Graduate Examination

To graduate with a MS in Health Data Science, students must pass the Masters Comprehensive Examination. That examination is administered after a student has successfully completed: (1) the required coursework including Categorical Data Analysis (HDSC 835), Linear Regression (HDSC 840), Statistical Learning I (HDSC 881), and Observational Health Data Analysis (HDSC 861); and (2) during the final semester of enrollment. The examination has two purposes: to assess the student's strengths and weaknesses and to determine whether the student should be awarded the MS degree. The examination is created and administered by a committee of at least three members of the Department Graduate Faculty. If this examination is failed, a second examination may be taken no sooner than three months later and is subject to committee approval. The committee can recommend that the student leave the program following the semester in which the examination is taken. After two failures, no further examination is permitted, and the student will not be awarded the MS degree.

VIII. Core Faculty

Note: * Next to Faculty Name Denotes Director of the Program, if applicable
FTE: 1.0 FTE = Full-Time Equivalency Devoted to Program

Faculty Name	Rank	Highest Degree	Tenure Track Y/N	Academic Area of Specialization	FTE to Proposed Program
Jonathan Mahnken	Professor	PhD	Y	Biostatistics	0.10

Jianghua He	Associate Professor	PhD	Y	Biostatistics	0.10
John Keighley	Associate Professor	PhD	N	Biostatistics, Data Science	0.20
Jo Wick	Associate Professor	PhD	Y	Biostatistics	0.15
Dong Pei	Research Instructor	PhD	N	Bioinformatics, Informatics, Data Science	0.10
Jinxiang Hu	Assistant Professor	PhD	Y	Biostatistics, Data Science	0.20
Devin Koestler	Associate Professor	PhD	Y	Biostatistics, Bioinformatics, Data Science	0.00
Jeffrey Thompson	Assistant Professor	PhD	Y	Biostatistics, Bioinformatics, Informatics, Data Science	0.00
Dinesh Pal Mudaranthakam	Teaching Associate	MS	N	Informatics, Data Science	0.10
Lynn Chollet Hinton	Assistant Professor	PhD	Y	Bioinformatics, Informatics, Data Science	0.15
Prabhakar Chalise	Associate Professor	PhD	Y	Biostatistics, Bioinformatics, Data Science	0.15
Yanming Li	Assistant Professor	PhD	Y	Biostatistics, Bioinformatics, Data Science	0.05
Mihaela Sardiu	Associate Professor	PhD	Y	Bioinformatics, Data Science	0.20
Milind Phadnis	Associate Professor	PhD	Y	Biostatistics	0.10

Number of graduate assistants assigned to this program **0**

The Department faculty who are responsible for teaching courses in this program include:

Jonathan Mahnken, PhD. Professor with Tenure. Research focus: study design; power and sample size calculations; secondary data analysis; survival analysis; categorical data analysis; statistical consulting and expert testimony; analysis of health claims data. Dr. Mahnken has an extensive collaboration record that includes basic science and translational research in stroke, clinical trials, and health services research. He is the director of the Data Management and Statistics Core of the NIA-funded KU Alzheimers's Disease Center and was PI of an R03 from the National Institute of Dental and Craniofacial Research of the NIH. Dr. Mahnken currently teaches BIOS/STAT 805: Professionalism, Ethics and Leadership in the Statistical Sciences.

Jianghua He, PhD. Associate Professor with Tenure. Research focus: survival analysis; time varying coefficient models; meta-analysis; missing data; Bayesian dynamic survival models. Dr. He has developed statistical methodology for problems in obesity and nursing research, specifically in the area of survival analysis. She currently teaches BIOS 872: Mathematical Statistics II, and BIOS 835: Categorical Data Analysis.

John Keighley, PhD. Education Associate Professor, Education Track. Research focus: Analysis and Reporting of Cancer Registry Data. Dr. Keighley has worked with the Kansas Cancer Registry and on various other cancer- related grants and studies. He has worked with researchers that specialize in engineering, smoking cessation, genitourinary disease and gastroenterology, and large data sets such as Kansas Medicaid, and the SEER-Medicare linked database. He currently teaches BIOS 820: SAS programming I, BIOS 850: Multivariate Statistics, and BIOS 880: Data Mining and Analytics.

Jo Wick, PhD. Associate Professor with Tenure/Director of Graduate Education. Research focus: Statistics education, Bayesian statistics and applications, clinical trial design. Dr. Wick has extensive experience in clinical trial design and has worked with researchers in cardiology, nursing, endocrinology, cancer, and early phase drug discovery. Her primary role within the department is to oversee the Graduate Programs in Biostatistics and Applied Statistics (MS/PhD) where she has received the Outstanding Graduate Teaching Award, voted on by students. Her research is focused on Bayesian statistics and data analysis and statistics education. She currently teaches BIOS/STAT 840: Linear regression and has developed the new course STAT 805: Professionalism, Leadership and Ethics for Data Scientists. She has previously taught BIOS 830: Experimental Design.

Dong Pei, PhD. Research Instructor and Bioinformatics Specialist. Research Focus: Development and application of bioinformatics tools/pipelines to analyze high-throughput 'Omic data, including: DNAseq, RNAseq, scRNAseq, and Infinium Methylation microarray. Dr. Pei currently teaches STAT/DATA 819: Introduction to Python, BIOS/STAT 823: Introduction to Programming and Applied Statistics in R.

Jinxiang Hu, PhD. Assistant Professor, Tenure track. Research Focus: Patient reported outcome, health disparity, structural equation modeling, item response theory, longitudinal modeling, mixture modeling, machine learning. Dr. Hu currently teaches BIOS/STAT 833: Measurement for Statisticians, STAT/DATA 881: Statistical Learning I, and STAT/DATA 882: Statistical Learning II.

Devin Koestler, PhD. Associate Professor with Tenure/Associate Director Biostatistics and Informatics Shared Resources. Research focus: high-dimensional genomic data, statistical genomics, mixture models, clustering and classification, molecular epidemiology, epigenetics, and DNA methylation. Dr. Koestler's research focus is the development and application of statistical methods for high-throughput 'omics' data; in particular, array-based DNA methylation data. Dr. Koestler has previously taught BIOS/STAT 830: Experimental Design.

Jeffrey Thompson, PhD. Assistant Professor, Tenure track. Research focus: Statistical/Machine learning methods, data integration, feature selection, quantitative 'omics, molecular epidemiology, survival analysis, and predictive models. Dr. Thompson developed and taught three courses STAT/DATA 824: Data Visualization and Acquisition, STAT/DATA 881: Statistical Learning I and STAT/DATA 882: Statistical Learning II.

Dinesh Pal Mudaranthakam, MS. Teaching Associate. Research focus and technical support: Design and implementation of Data warehouse such as Cancer Curated Clinical Outcomes Database, Research Databases to collect clinical research information, Query Module to validate study design feasibility, Software module to identify early cancer patient, Clinical Integration (Epic and EResearch), Investigator Initiated trial study build, standard eCRF (electronic case report form), Data dissemination for analysis and administrative purpose, patient accrual tracking and Managing the Biospecimen Inventory Software (OpenSpecimen) for the University of Kansas Cancer Center (KUCC). Mr. Mudaranthakam currently teaches DATA 822: Introduction to SQL, and DATA 817: Introduction to Tableau.

Lynn Chollet Hinton, PhD. Assistant Professor, Tenure track. Research focus: Population health, epidemiology, biomarkers, electronic health records, administrative claims data, observational study design, missing data, categorical data analysis, survival analysis, longitudinal data, and data integration. Dr. Hinton is currently developing new course specifically designed for proposed degree MS in Health Data Science, DATA 861: Observational Health Data Analysis.

Prabhakar Chalise, PhD. Associate Professor with Tenure/Assistant Director of Graduate Education. Research focus: Statistical Genetics and Genomics, Computational Statistics, Survival Analysis, Methods in Biostatistics. Dr. Chalise's primary research interest is in the development and application of statistical methods to health sciences research. Dr. Chalise's educational leadership role includes overseeing the Masters and PhD education program in Biostatistics. Dr. Chalise currently teaches BIOS 871: mathematical Statistics I, and BIOS 855: Statistical Methods in Genomics Research.

Yanming Li, PhD. Assistant Professor, Tenure track. Research focus: High-dimensional Data Analysis; Variable Selection; Survival Analysis with High-Dimensional Predictors; Weak Signal Detection, Estimation and Their Effects in Prediction; Probabilistic Graphical Models; Computational Statistics; Cancer Genomics; Neuroimaging-Genomics. Dr. Li is currently developing contents for course DATA 818: Data Summarization and Management.

Mihaela Sardiu, PhD. Associate Professor, Tenure track. Research focus: Dr. Sardiu's research interest is in the field of quantitative omics data, with a focus on the development of computational methods for processing and extracting biological information from large and complex datasets. Another research interest of Dr. Sardiu is to provide a holistic view of merged data and interpretation tools that harmonize biological information across heterogeneous platforms. Dr. Sardiu is currently teaching DATA 824: Data Visualization and Acquisition.

Milind Phadnis, PhD. Associate Professor with Tenure. Research focus: Dr. Phadnis's research interest is in the field of Survival Analysis and Design of Clinical Trial. He has collaborated extensively with researchers from nephrology, stroke, neurology, oncology, and many other areas. He is the co-director of the Biostatistics section of Investigator Initiated Trials supported by the KU Cancer Center. Dr. Phadnis is currently teaching BIOS 845: Survival Analysis.

Other faculty involved in curriculum, admissions, and future teaching include:

Matthew Mayo, PhD. Chair and Professor with Tenure. Research focus: robust regression; linear models; experimental design. Dr. Mayo currently teaches advanced courses within the MS/PhD programs in Biostatistics. His research focus is in clinical trial design and robust methods for regression.

Byron Gajewski, PhD. Professor with Tenure. Research focus: Bayesian data analysis; latent variable modeling. Dr. Gajewski currently teaches advanced courses within the MS/PhD programs in Biostatistics. His research focus is Bayesian data analysis with applications to nursing and health care evaluation data.

No new faculty is required for the proposed program. Department has recently recruited a couple of new faculty of which some effort will be allocated to this program. This program consists of all but one course already offered through the Department of Biostatistics and Data Science and taught by Department faculty (listed above). There will be one new course unique to this program which will be developed by the subject area specific experts from the current faculty in the Department.

IX. Expenditure and Funding Sources *(List amounts in dollars. Provide explanations as necessary.)*

A. EXPENDITURES	First FY	Second FY	Third FY
Personnel – Reassigned or Existing Positions			
Faculty (*for one new course*)	$84,642	$87,181	$87,181
Administrators *(other than instruction time)*	$0	$0	$0
Graduate Assistants	$36,540	$37,636	$38,765
Support Staff for Administration (*e.g., secretarial*)	$0	$0	$0
Fringe Benefits *(total for all groups)*	$23,469	$23,938	$23,938
Other Personnel Costs	$0	$0	$0
Total Existing Personnel Costs – Reassigned or Existing *(All the cost for faculty and GTA have been covered already with existing state funds and tuition dollars. Therefore, the cost shown here is not the new cost.)*	$144,651	$148,755	$149,884
Personnel – New Positions			
Faculty	$0	$0	$0
Administrators *(other than instruction time)*	$0	$0	$0
Graduate Assistants	$0	$0	$0
Support Staff for Administration (0.5 *FTE*)	$32,500	$33,475	$33,475
Fringe Benefits *(total for all groups)*	$13,893	$14,019	$14,019
Other Personnel Costs	$0	$0	$0
Total Existing Personnel Costs – New Positions	$46,393	$47,494	$47,494
Start-up Costs - One-Time Expenses			
Library/learning resources	$0	$0	$0
Equipment/Technology	$0	$0	$0
Physical Facilities: Construction or Renovation	$0	$0	$0
Other	$0	$0	$0
Total Start-up Costs *(There will be no new cost. All the existing resources will be utilized.)*	$0	$0	$0
Operating Costs – Recurring Expenses			
Supplies/Expenses	$0	$0	$0
Library/learning resources	$0	$0	$0
Equipment/Technology	$2,000	$3,000	$3,000
Travel	$0	$0	$0
Other	$0	$0	$0
Total Operating Costs *(All the existing resources will be utilized except one possible new software.)*	$2,000	$3,000	$3,000
GRAND TOTAL COSTS	$48,393	$50,494	$50,494

B. FUNDING SOURCES *(projected as appropriate)*	Current	First FY (New)	Second FY (New)	Third FY (New)
Tuition / State Funds		$74,878	$149,756	$149,756
Student Fees		$16,940	$33,880	$33,880
Other Sources				
GRAND TOTAL FUNDING		$91,818	$183,636	$183,636
C. Projected Surplus/Deficit (+/-) (Grand Total Funding *minus* Grand Total Costs)		$43,425	$133,141	$133,141

X. Expenditures and Funding Sources Explanations

A. Expenditures
Personnel – Reassigned or Existing Positions

All but one of the program courses already exist, requiring no additional faculty effort for instruction. Any faculty effort indicated in Section IX is already accounted for and supported by existing graduate programs in Biostatistics. This is all paid for by the existing state funds and tuition dollars that are allocated to the Department through the SOM funding model.

Personnel – New Positions

One new Academic Program Specialist will be hired and devote 0.5 FTE towards the day-to-day management of the program and students.

Start-up Costs – One-Time Expenses

None

Operating Costs – Recurring Expenses

Software purchase for the proposed new course.

B. Revenue: Funding Sources

The funding sources will be tuition and fees as shown in section IX above. It was assumed that there will be 80% in-state and 20% out-of-state students. The average tuition per student will be $421.15 × 0.8 + $989.6 × 0.2 = $534.84. For the first year, average credit hours for each student will be 14 (19 Credits for full-time and 9 Credits for part-time). Therefore, total tuition revenue from 10 students in first year will be $534.84 × 14 × 10 = $74,878. Course fee for 14 average credit hours for 10 students at the rate of $121 will be $16,940. Therefore, the total revenue including tuition and student fees will be $74,878 + $16,940 = $91,818. In second year, we will have 10 new students (5 full time, and 5 part time) resulting in 20 students (10 full time and 10 part time). This will double the revenue to $183,636. After second year, at least 10 students will be admitted and around 10 students will graduate per year resulting in 20 (10 full time and 10 part time) students at any given year. Therefore, the tuition and fee revenue after second year will be at least $183,636.

C. Projected Surplus/Deficit

In the first year, there will be an expected surplus of $43,425. After first year, we expect a surplus of at least $133,141 per year.

XI. References

Davenport T.H. and Patil D.J. Data scientist: The sexiest job of the 21st century. (2012, October 1). Harvard Business Review. https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century

Best jobs in America. (n.d.). Glassdoor. https://www.glassdoor.com/List/Best-Jobs-in-America-LST_KQ0,20.htm

Bureau of Labor Statistics (2020). https://www.bls.gov/ooh/fastest-growing.htm

August LinkedIn workforce report: Data science skills are in high demand across industries. (2018). LinkedIn. https://news.linkedin.com/2018/8/linkedin-workforce-report-august-2018

Flowers A. Data scientist: A hot job that pays well. (2021, February 9). Indeed Hiring Lab. https://www.hiringlab.org/2019/01/17/data-scientist-job-outlook/

Workforce development - MARC data and economy section. (2020). MARC - Mid-America Regional Council - Regional Planning for Greater Kansas City. https://www.marc.org/Data-Economy/Workforce-Development.html

Sheng, J., Amankwah-Amoah, J., Khan, Z., & Wang, X. (2020). COVID-19 pandemic in the new era of big data analytics: Methodological innovations and future research directions. *British Journal of Management*. https://doi.org/10.1111/1467-8551.12441