Proposal to Establish a Self-Supporting Graduate Professional Degree Program
Master of Data Science and Analytics at UC Merced

Original Submission - October 13, 2020

First Revision - December 14, 2020

Second Revision – September 1, 2022

Contacts

Christian Fons-Rosen
Associate Professor of Economics
cfons-rosen@ucmerced.edu

Alex Petersen
Associate Professor of Management
apetersen3@ucmerced.edu

Lace Padilla
Assistant Professor of Cognitive Science
Lace.padilla@ucmerced.edu

Proposal Team

John Abatzoglou, Professor of Management

Christian Fons-Rosen, Associate Professor of Economics

Rowena Gray, Associate Professor of Economics

Hanna Gunn, Assistant Professor of Philosophy

Dan Hicks, Assistant Professor of Philosophy

Andrew Johnston, Assistant Professor of Economics

Crystal Kolden, Assistant Professor of Management

Alvin Cha, Assistant Director, Division of Management and Information

Paul Maglio, Professor of Management

Lace Padilla, Assistant Professor of Cognitive Science

Alex Petersen, Associate Professor of Management

Table of Contents

Executive Summary

We propose establishing a twelve-month intensive *Master of Data Science and Analytics*, a terminal self-supporting professional graduate degree program to be administered by the proposed E & J Gallo School of Management at the University of California, Merced. The program will be run jointly by three existing Graduate Groups – *Cognitive and Information Sciences, Economics,* and *Management of Innovation, Sustainability, and Technology* – with faculty oversight and a program director from these groups.

What does the program entail?

The program will provide a core curriculum designed around practical skills that successful students can apply in diverse information-driven public and private industry settings. Major themes addressed by the program include practical economic principles underlying the data revolution, principles and best practices for effective communication with data, analytical skills for disclosing the embedded values and norms in data science and computing, introduction to scripting languages for data management and visualization, design of interactive web- and smart-phone-based platforms to facilitate organizational systems thinking and real-time decision-making, and best practices and management tools for data-intensive team-work. A summer capstone project will provide students with an integrative team-oriented experience to prepare for time-crunch, logistical, and other operational constraints encountered in real world decision-making scenarios.

Why this program?

There is increasing public and private demand for the development of a resilient US workforce in data science and analytics, geographic information systems, interactive information-system design, and data-driven policy making. This program is designed to be attractive to prospective students seeking to enter this informatics workforce segment, which according to current national and international trends is likely to continue growing. To address this educational need, the proposed core curriculum will develop students' *soft* computational skillset – one oriented around data communication, inferential analytics, ethics, and visualization – tailored for real-world applications, inspired by elective coursework and implemented in the capstone project. This approach differs from the regimen found in more traditional computer science-based data science masters programs, which also emphasize theoretical aspects of hardware, software and their integration, and hence require undergraduate training in or around computer science, with normative graduation times of at least two years.

By adopting a one-year cohort model that leverages interdisciplinary methods and scope, students of the proposed program will receive integrated training in a high-demand skillset that will increase their competitiveness for information technology and data analytics jobs. Enrollment will not be contingent on prior educational background (i.e., no prior coding or technical experience will be required, unlike many other professional data science programs). Consequently, students from different backgrounds will be immersed in multi-disciplinary teams that emulate real-world settings where effective communication and hybrid approaches to problem-solving are critical to success. The location of UC Merced provides students with a cost-effective alternative to similar in-person programs offered in locations with high cost-of-living burden, while being close to emerging data analytics hubs in the state. The advantage of

this program over related online programs is the intensity (one year), relatively high rigor-to-cost ratio, and a curriculum designed around integrative team-oriented practical problem-solving. As such, this program will focus on the entry-level skillset demanded by both traditional analyst positions in private and public organizations, and also emergent data-oriented job sectors.

How will the program operate?

An existing group of faculty from the proposed E & J Gallo School of Management will deliver the core curriculum through a combination of on-load and off-load teaching supported by program revenues. Program tuition will be set at the low end of similar UC programs, and at least 20% of program revenue will be returned to student aid. As such, this program will appeal to students who desire a non-PhD track program that is cost-effective yet comprehensive, interactive yet personalizable, and practical rather than theoretical. Students will take a small set of core classes, with options to specialize through electives focused on sustainability and environment, human behavior, or policy and decision-making. By design, this program can be implemented using existing campus resources and leveraging the cost-efficiency of computational resources – personal laptops, internet access, and open data sources – which are all relatively inexpensive, diversifiable, and reusable.

Introduction

We propose to establish a novel twelve-month self-supporting professional graduate degree program, a *Master of Data Science and Analytics (MDSA)*, to be overseen jointly by three existing Graduate Groups – *Cognitive and Information Sciences*, *Economics*, and *Management of Innovation, Sustainability, and Technology*. The proposed new program would be administered by the proposed E & J Gallo School of Management at the University of California, Merced. The proposed data science program would be unique in the UC system, integrating foundational informatics practicum with essential soft skills for data communication, open data and code-sharing frameworks for leveraging team science, ethical use of data in visualizations and applications, and incorporating tracks in data analytics related to sustainability and environment, human behavior, and policy and decision making. The proposed program would be the first self-supporting graduate professional degree program at UC Merced.

Aims and Objectives

The demand for highly trained computational and data scientists is growing at an astounding rate. To fully realize the promise of "big data" to foster innovation – improving management of critical challenges in environmental sustainability, human health, living standards, national security and much more – it is not enough to just increase computing power and data storage capacities. We must increase skills related to data, information, decision making and their application to management. Data sets are increasing in complexity just as fast as they are increasing in size. Relationships among variables are often nonlinear and context-dependent, and useful information must be culled from virtually limitless sources of data. Decisions often must be made under conditions of time pressure and uncertainty, and systems must constantly adapt to new data and changing conditions and management objectives. Data scientists bridge the gap between data and computation on the one hand, and information and decisions on the other. Such *human-in-the-loop* computing has become a dominant paradigm in which the inputs and outputs of predictive analytics are not the end-all. Rather, another essential piece of the data pipeline are expert data science communicators, needed to translate the added value of data insights into real-world action. Big data is creating job opportunities in many fields related to computer science and math: The Bureau of Labor Statistics projects employment of computer and information scientists will grow 16% from 2018-2028, more than three times faster the national growth rate for all occupations.[1] But the impact of the data science revolution and the opportunities it creates will also be realized in careers far afield from computer and information sciences alone.

The MDSA degree program will equip students to draw sound conclusions from data in context, using principles of statistical inference, computational processes, geographic information systems, data management strategies, domain knowledge, ethical analysis and probability and statistics theory. Students will learn to carry out analyses of data through the full cycle of investigative processes in scientific and managerial contexts. They will gain a deep appreciation of the human, social, and institutional structures and practices that shape technical work around computing and data, as well as an understanding of how data, data analytics, artificial intelligence, and computing permeate and shape our individual and social lives. This approach differs from the regimen found in more traditional computer science-based data science masters

[1] See https://www.bls.gov/ooh/computer-and-information-technology/computer-and-information-research-scientists.htm#tab-1

programs, which also emphasize theoretical aspects of hardware, software and their integration, and hence require undergraduate training in or around computer science, with normative graduation times of at least two years. By adopting a one-year cohort model that leverages interdisciplinary methods and scope, students of the proposed program will receive integrated training in a high-demand skillset that will increase their competitiveness for information technology and data analytics jobs.

Thus, the masters is envisioned to be a one-year (two semesters plus one summer) self-supporting graduate professional degree program (SSGPDP). Coursework will include an introduction to data science and data science methods courses, as well as courses in ethics, statistics, and computational modeling, with electives organized into themes – *Sustainability and Environment, Human Behavior, Policy and Decision Making* – drawn from existing graduate programs, including *Cognitive and Information Sciences (CIS), Economics (ECON), Environmental Systems (ES), Psychological Sciences (PSY), Public Health (PH), Political Science, Management of Innovation Sustainability and Technology (MIST),* and more. The program will also include a capstone activity in the summer, either as an internship or a research practicum.

Background and Historical Development

Recent research and technological advances are changing how we conceptualize and interact with data science. From viruses and amoebae to swarms and human brains, scientists are making rapid advances into the data sciences that allow us to uncover the mechanisms and principles that underlie an increasingly remarkable range of natural phenomena. At the same time, data science is becoming just as pervasive in driving technological change. From devices and software that anticipate customer behaviors to robots and algorithms that autonomously seek out information in real and virtual spaces, data science is fast becoming an integral part of modern life. And data science is increasingly used by social scientists to understand the statistical and causal regularities that drive emergent phenomena in economic and social systems, leading to new insights into the impacts of policies and events.

Researchers from various disciplines have begun to address these fundamental questions, including mathematicians, physicists, biologists, environmental scientists, geographers, computer scientists, economists, philosophers, and behavioral and cognitive scientists. Common to the studies of all these scientists are basic principles of *complex systems*. Intelligent functions always emerge from interactions among many system components that can adapt in ways that give rise to many different patterns of behavior depending on conditions. Intelligent functions are often enhanced, even optimized, when component interactions are balanced between extremes, for example, order versus disorder, independence versus dependence, excitation versus inhibition, and so on. Intelligent functions are learned and shaped on the basis of relatively small, incremental changes to individual components, yet learning can be swift and based on very sparse information. Taming this complexity by way of multi-disciplinary teams that harness hybrid human-computer capabilities is both a motivating paradigm and design principle underlying the proposed program.

Another consideration is that both theoretical and practical advances are likely to emerge from academia and will complement important technological advances in data science that arise from industry. Indeed, the big data revolution, along with the imminent age of robotics, is likely to

drive major technological innovations and advances. Many such advances will be peripheral to this proposal, for instance, in software and hardware needed to process vast amounts of data and allow robots to operate robustly for extended periods and in various environments. However, advances also will be needed to build systems that can collect, manage, and act upon information to manage complex systems for beneficial goals, and stimulating these advances will be a core objective of our data science programs. For instance, companies will need systems that adapt to and predict the needs and specifications of users and clients, or that can intelligently filter and highlight data that are autonomously identified as high priority. Other companies will need systems that interact with the physical world in semi- or fully autonomous modes to manage transport or sensor systems, for example, or search large environments that are too dangerous or expensive for humans or manage processes that are hazardous to humans. In the policy realm, the analysis of large datasets will allow policymakers to harness new methods of causal inference to understand the key determinants and scalar dependencies of policy outcomes. These examples demonstrate the need for applied scientists with broad perspectives who can perceive opportunities and problems, and formulate strategies and solutions, that require synthetic knowledge of data science.

Proposed Timeline

The *Master of Data Science and Analytics* is proposed to begin in Fall 2024, aligned with the start of the proposed E & J Gallo School of Management at UC Merced (proposed separately). Enrollments are projected to start at 10 students and grow to a steady-state of 30 students over five years (see Table 1). As described below, we believe we can accommodate these students in existing graduate courses.

TABLE 1: PROJECTED MDSA ENROLLMENTS AT UC MERCED

	AY 2024	AY 2025	AY 2026	AY 2027	AY 2028
MDSA	10	15	20	30	30

Relationship to Campus Academic Plan and Other Programs

The UC Merced Campus has engaged in strategic academic planning process to align individual department and school plans and priorities with campus priorities.[2] Overall campus priorities focus on creating UC-quality research and educational programs and boosting diversity of programs and of students and faculty. We believe that the proposed MDSA program aligns with campus priorities by:

1. Establishing an interdisciplinary graduate program that increases the diversity of offerings and enrollments at the graduate level.
2. Increasing enrollments in graduate courses across a number of programs, which in turns increases demand and access to valuable educational programs.
3. Generating revenues for individual graduate groups and departments to support core academic operations, particularly PhD student research and teaching, potentially increasing research expenditures and productivity, as well as graduate student success.

[2] https://strategicplan.ucmerced.edu/

4. Paying careful attention to issues of equity, diversity and inclusion for faculty and students, as described below.

Data Science and Analytics is one unifying theme of the proposed Gallo School at UC Merced, whose vision it is to integrate across the traditional disciplines found in the existing three UC Merced schools (Natural Sciences; Engineering; Social Science, Humanities & Arts). Such integration would give rise to diverse new educational pathways, thereby providing both undergraduate and graduate students an even richer – and potentially transformative – developmental experience, leading to novel postgraduate career trajectories.

The most closely related programs on campus are in Applied Mathematics and in Computer Science and Engineering. In particular, the B.S. in Applied Mathematics with emphasis in Computational and Data Sciences adds courses in differential equations, linear algebra, optimization, statistics, and modeling stochastic processes to the core foundation of a major in mathematics; and the B.S. in Computer Science and Engineering may incorporate electives in database systems and machine learning. Other programs also have courses related to data science, particularly statistical methods applied to specific application domains, including Biology, Environmental Engineering, Public Health, Political Science, and Psychology. In addition, the Library offers a variety of services related to data science analytics, including workshops and programs related to data literacy and data carpentry as well as the Spatial Analysis and Research Center (SpARC). Our aim is to make use of existing resources wherever possible and practical, and to incentivize other programs appropriately, for instance, providing increased enrollments and resources to help improve programs and course offerings.

The proposed MDSA program complements existing research centers and undergraduate and graduate programs at UC Merced. To begin, the proposed M.S. program is designed around three thematic course tracks (Sustainability and Environment, Human Behavior, and Policy and Decision-making) that are well-aligned with the campus' informatics-oriented research organizations:

- Center for Information Research in the Interest of Society (CITRIS)
- Health Sciences Research Institute (HSRI)
- Sierra Nevada Research Institute (SNRI)
- Spatial Analysis and Research Center (SpARC)
- Data Science Institute (DSI)

Moreover, because the proposed curriculum converges around an integrative capstone project that can be overseen by any faculty mentor on campus (as well as industry, government, or other non-profit partners), this program could facilitate research support for short-term real-world projects by matching students with ongoing projects in these organizations.

Two existing undergraduate campus activities that align with the proposed program are the *Mobile App Challenge* and *HackMerced*, which attract more than ~100 and ~300 students per year, respectively. These events would serve as recruitment opportunities for the proposed M.S. program, which could provide a natural continuation of the applied informatics pathway for UC Merced undergraduates, as many participants find themselves increasingly inclined towards post-

graduate educational opportunities in data science and analytics, given its lucrative yet resilient job market.[3]

Indeed, many UC Merced undergraduates are eager to take part in extracurricular data science activities on campus. For example, *HackMerced* is a 36-hour student-organized weekend hackathon that draws participants from campus and as far as the Bay Area and is officially sponsored by *Major League Hacking*. This enthusiasm in part owes to the rural location and surroundings of the Merced campus. As such, students naturally gravitate towards convenient and affordable on-campus entertainment options, and computer-oriented activities that require just a laptop and internet access are appealing. Hence, it is likely that a one-year graduate program that offers the opportunity to develop, integrate, and enhance one's data analytics skills would appeal greatly to the burgeoning demand for data science and analytics experiences that already exist at the undergraduate level on campus.

We expect that at least to start, the main undergraduate pipeline will flow from UC Merced graduates from degree programs that are quantitative, data- and computation-oriented, such as:

- Applied Mathematical Sciences, B.S.
  - Computational and Data Sciences Emphasis, B.S.
  - Applied Mathematical Sciences, Computer Science Emphasis, B.S.
  - Applied Mathematical Sciences, Economics Emphasis, B.S.
  - Applied Mathematical Sciences, Environmental Emphasis, B.S.
- Environmental Systems Science, B.S.
- Economics, B.A. & B.S
  - Economics, B.A.
  - Economics, Economic Analysis and Policy Emphasis, B.S.
  - Economics, Quantitative Economics Emphasis, B.S.
- Cognitive Science, B.A & B.S
  - Cognitive Science, B.A.
  - Cognitive Science, B.S.
- Computer Science and Engineering, B.S.
- Management and Business Economics, B.S.
- Management Analytics and Decision-Making, Minor
- Physics, B.S.
- Public Health, B.S.

These interdisciplinary programs may serve as the primary enrollment feeders for MDSA providing disciplinary breadth that contributes to the interdisciplinary program – by design.

The proposed program also builds around existing graduate-level initiatives, such as the *2018 Data Science Summit* sponsored by UCM Graduate Division, and the *UC Merced Data Science Challenge* in partnership with Lawrence Livermore National Lab. UC Merced is also host to an *NSF National Research Training - Interdisciplinary Computational Graduate Education Program* (NRT-ICGE), which has focused on providing incoming graduate students with computational and data analytics skills to enhance success and reduce attrition rates. Other campus initiatives, such as the *NSF AGEP California HIS Alliance* and *NSF INCLUDES and*

[3] Data from Glassdoor, Inc. cited in: Say goodbye to six-figure starting salaries – with these exceptions, CNBC. (6/8/2020)

Women in Science, Technology, Engineering, and Math (W-STEM) programs, are particularly oriented for underrepresented minorities and first-generation students, identifying the campus' role as a designated Hispanic-Serving Institution to develop STEM pathways that provide diversity-oriented value-added to the workforce,[4] which increasingly calls for computational skillset.[5] Such a skillset is in high demand – both nationally and regionally, given the proximity to CA hubs: Los Angeles (media and entertainment, trade, fashion), San Diego (pharmaceuticals) and San Francisco (information technology, high-performance computing, web and social networking services) – as well as by agriculture,[6] the dominant industry in the Central Valley.

The proposed MDSA program also builds on UC Merced Library resources and services. With the Library, we are developing a "Data Bootcamp" for students requiring background training in data science tools, techniques, and languages – based on the Library's existing partnership with The Carpentries and data literacy workshops and resources for Data Curation and Scholarship. In addition, we will rely on Library resources for GIS, including workshops and software, provided by SpARC.

At the graduate education level, the proposed masters program leverages existing classes and existing ladder-rank FTE. Only two new core courses would be required to launch the proposed program: *Interactive Data Development* and the *Capstone Project*. The remaining electives are drawn from existing graduate programs on campus.

There are only a few masters programs on campus that are empirical and thus data-oriented in nature:

- Applied Mathematics, M.S.
- Cognitive and Information Sciences, M.S.[7]
- Economics, M.A.
- Management of Complex Systems, M.S.
- Public Health, M.S.P.H.
- Quantitative and Systems Biology, M.S.

However, these programs are primarily milestone degrees or off-ramps for students initially pursuing a traditional research and discipline-oriented Ph.D. Hence, given the one-year, integrative, interdisciplinary, and applied orientation of the proposed program, we do not

[4] National Academies of Sciences, Engineering, and Medicine, 2019. Minority serving institutions: America's underutilized resource for strengthening the STEM workforce. National Academies Press.

[5] National Academies of Sciences, Engineering, and Medicine, 2017. A 21st century cyber-physical systems education. National Academies Press; National Research Council, 2001. Building a workforce for the information economy. National Academies Press; National Academies of Sciences, Engineering, and Medicine, 2017. Building America's skilled technical workforce. National Academies Press; National Academies of Sciences, Engineering, and Medicine, 2017. Information technology and the US Workforce: Where are we and where do we go from here? National Academies Press; National Research Council, 2015. Preparing the workforce for digital curation. National Academies Press; National Academies of Sciences, Engineering, and Medicine, 2016. Promising practices for strengthening the regional STEM workforce development ecosystem. National Academies Press; National Academies of Sciences, Engineering, and Medicine, 2017. Strengthening Data Science Methods for Department of Defense Personnel and Readiness Missions. National Academies Press.

[6] National Academies of Sciences, Engineering, and Medicine, 2019. Improving Data Collection and Measurement of Complex Farms. National Academies Press.

[7] The new Cognitive and Information Sciences M.S. program specifically targets students interested in the M.S. track only.

anticipate it having a significant impact on enrollment in the Ph.D. programs that also offer these masters degrees.

The course-level resource requirements for the proposed MDSA program are primarily computational, and thus localized to each student's personal computer. Thus, we do not foresee adverse negative resource impacts on other existing graduate groups. The only other campus master's program that students might consider as an alternative to the MDSA program is the graduate program in Electrical Engineering and Computer Science (EECS), which offers both Thesis and Non-thesis MS options. To mitigate confusion among potential students, MDSA recruitment information will explicitly identify the fundamental differences in scope regarding methods, theory and applications between a traditional EECS graduate program and this interdisciplinary informatics program. One metric that the MDSA program will follow to understand and possibly avoid confusion is to track the number of applications to both programs over time – in principle, assuming that students have a clear idea of the aims of these two different programs, we expect a relatively small number of co-applications.

Relationship to Programs at other UC Campuses

Data science and business analytics programs have multiplied rapidly across the system over the last few years, with related masters programs at seven of the ten UC campuses (see Table 2). Some programs are run through engineering schools and focus on data science, others are run through management schools and focus on business analytics, and yet others are online programs that focus on data science.

TABLE 2: RELATED PROGRAMS IN THE UC SYSTEM

Data Science Programs		
Degree Title	**Campus**	**School**
Master of Environmental Data Science	Santa Barbara	Environmental Science and Management
Master of Engineering - Data Science and Systems Concentration	Berkeley	Engineering
Master of Science in ECE - Machine Learning and Data Science Focus	San Diego	Engineering
Master of Advanced Study in Data Science and Engineering	San Diego	Engineering
Business Analytics Programs		
Degree Title	**Campus**	**School**
Master of Science in Business Analytics	Davis	Management
Master of Science in Business Analytics	Los Angeles	Management
Master of Science in Business Analytics	San Diego	Management
Master of Science in Business Analytics	Irvine	Management
Online Programs		
Degree Title	**Campus**	**School**
Master of Information and Data Science	Berkeley	Information
Master of Science in Engineering with Certificate of Specialization in Data Science Engineering	Los Angeles	Engineering
Master of Science in Engineering - Specialization in Data Science	Riverside	Engineering

Data Science Programs

Of the programs we surveyed, *UC Santa Barbara's Master of Environmental Data Science*[8] is the most similar to our proposed program because of its focus on an application area, in this case, the analysis of environmental data; it is an 11-month professional degree in data science. This program is held on-campus and students must be enrolled full-time. Core faculty come from the Bren School of Environmental Science and the National Center for Ecological Analysis and Synthesis, with some additional support from UCSB's Department of Geography and Department of Economics. Because key faculty specialize in addressing environmental and ecological research questions, courses are focused on environmental matters – from "EDS 223 Spatial Analysis for Environmental Data Science" in the Fall to "EDS 232 Machine Learning in Environmental Science" in the Winter and to "EDS 230 Modeling Environmental Systems" in the Spring – the program revolves entirely around environmental sciences.

In our case, we aim to exploit the synergies of the three of the four graduate groups involved in the proposed Gallo School (Cognitive and Information Sciences, Economics, Management of Innovation, Sustainability, and Technology). Rather than focusing fully on a single area, our program aims to combine our strengths in data-driven research across major disciplines to provide a set of fundamental skills applicable across disciplines. Our interdisciplinary program extends beyond the core expertise of faculty in the three graduate groups of the proposed Gallo School. For the next generation of data scientists, it will be common practice to have, for example, biologists working with computational scientists and engineers on data science problems; interdisciplinarity is embedded in the core of our values and courses offered. To this end, in addition to our core courses we have three branches of specialization: (i) Sustainability and Environment; (ii) Behavior; (iii) Policy and Decision Making. Obviously, one of our themes has a certain amount of overlap with the *Master of Environmental Data Science*, but two do not, and thus broaden the spectrum of students to be recruited and skills to be taught.

Other data science programs include, (1) *UC Berkeley's Master of Engineering - Data Science and Systems Concentration*,[9] which is housed in the Department of Electrical Engineering and Computer Sciences, requires students to have a strong technical background in computer science and engineering, and includes courses such as, Robotics and Embedded Software, and Signal Processing and Communications; (2) *UCSD's Master of Science in ECE - Machine Learning and Data Science Focus*[10] and (3) *UCSD's Master of Advanced Study in Data Science and Engineering*,[11] both of which are administered by the Jacobs School of Engineering and are targeted at mid-career engineering professionals with at least two years of work experience and an undergraduate degree in a highly quantitative field. Our proposed program differs substantially from all these, as ours does not focus on computer science or engineering topics nor require a technical education or professional experience.

[8] https://www.bren.ucsb.edu/admissions/master-environmental-data-science.htm

[9] https://eecs.berkeley.edu/academics/graduate/industry-programs/meng[9]

[10] http://www.ece.ucsd.edu/graduate/masters-of-science-ece-machine-learning-data-science-focus

[11] http://jacobsschool.ucsd.edu/mas/dse/

Business Analytics Programs

Five campuses offer masters degrees in *Business Analytics* through their business or management schools. These programs and their faculty focus squarely on data analytics applied to business data and business problems. *UC Davis's Master of Science in Business Analytics*[12] is a one-year program targeted at business professionals, with classes on weekends. *UCLA's Master of Science in Business Analytics*[13] is a 15-month program that emphasizes business needs and fundamental understanding of the economics of the firm; it includes a summer internship with a business partner. *UCSD's Master of Science in Business Analytics*[14] is an 11-month program focused on solving practical business problems and in using analytics to drive business decisions. *UCI's Master of Science in Business Analytics*[15] is a one-year program focused on use of data analytics in business contexts, such as marketing and operations. *UCR's Master of Science in Business Analytics*[16] is a 9-15-month program focused on data analysis and statistical computing to describe marketing, logistics, and financial patterns, forecasting revenues and consumer demand and identifying trends in financial markets using financial data. Our proposed program differs substantially from these, as ours does not focus on uses of data analytics in business settings nor require professional experience.

Online Programs

UC Berkeley's Online Masters of Information and Data Science,[17] *UCLA's Master of Science in Engineering with Certificate of Specialization in Data Science Engineering*,[18] and *UC Riverside's Master of Science in Engineering - Specialization in Data Science*[19] are online masters programs in areas of data science and analytics. Our proposed program is not intended to an online program, though we expect to incorporate some online and remote components over time.

Other Related Programs

There are many related programs in data science and analytics in California and across the country. For example, *University of San Francisco's Master of Science in Data Science (MSDS)*[20] is a full-time, one-year program with core courses that put special emphasis on linear regression and econometrics, machine learning, data structures and data systems; *USC's Master of Science in Environmental Data Science*[21] is a two-year program that combines courses in data science and sustainability and *USC's Master of Science in Public Policy Data Science*[22] is a one-to-two-year program for students with a quantitative background that focuses on decision making

---

[12] https://gsm.ucdavis.edu/msba-masters-science-business-analytics
[13] https://www.anderson.ucla.edu/degrees/master-of-science-in-business-analytics
[14] https://rady.ucsd.edu/programs/masters-programs/ms-in-business-analytics/
[15] https://merage.uci.edu/masters-business-analytics/
[16] https://business.ucr.edu/msba
[17] https://datascience.berkeley.edu/admissions/tuition-and-financial-aid/
[18] https://www.msol.ucla.edu/data-science-engineering/
[19] https://engineeringonline.ucr.edu/data-science/
[20] https://www.usfca.edu/arts-sciences/graduate-programs/data-science
[21] https://dornsife.usc.edu/environmental-studies/graduate-environmental-data-science/
[22] https://priceschool.usc.edu/programs/masters/public-policy-data-science

in specific policy contexts. Our proposed MDSA program is broader, incorporating aspects of each but with themes that cover topics in all of them.

Contributions to Diversity

The field of data science is not particularly diverse: A 2020 industry survey of 1,001 data scientists found that 71% were men.[23] A 2017 analysis of part-time students in general data science courses found that 46% were White and 28% were Asian, while only 8% Latinx and 4% African American.[24] The proposed MDSA program aims to help address this underrepresentation problem in data science as is designed to be open to students with broad undergraduate backgrounds, including the student population of UC Merced's undergraduate programs

A commitment to diversity in faculty, students, and research populations is an essential element of the current practices within each group and department collaborating on the MDSA program. In addition, MDSA will advance UC's goals for diversity, equity, and inclusion through:

- Recruitment and retention of underrepresented students, including LGBTQ, non-binary and disabled students, through funding support and outreach efforts, devoting at least 20% of program revenues for student aid, including fellowships for underrepresented students with demonstrated financial need.
- Integrating social justice and ethics into the curriculum, for instance through the core course on Data Ethics, and integration of ethical topics into other courses.
- Ensuring faculty from minoritized groups are in leadership roles, establishing a leadership pipeline from faculty across the constituent groups.
- Setting student tuition around $45,000, at the low end of comparable SSGPDPs at UC, including Berkeley ($70,000), Irvine ($50,000), and UCSD ($40,000),[25] helping to ensure access for a diverse student population.

More details about contributions to equity, diversity, and inclusion are provided in a separate section, *Equity, Diversity, and Inclusion* (pages 51-55).

Group that will Administer the Program

The MDSA Program will be overseen by a Program Director and an Executive Committee whose membership includes faculty from each of the three core graduate groups of the Program—CIS, ECON and MIST. By default, the Program Director serves as the Graduate Advisor for all program students but may assign other core program faculty to advise individual students as appropriate. The program will be administered by Dean of the proposed E & J Gallo School of Management. If the program operates prior to the establishment of the new Gallo School, the program will be administered by the Dean of the School of Engineering (and will move to the new school when possible).

[23] See https://365datascience.com/career-advice/career-guides/become-data-scientist-2020/

[24] See https://generalassemb.ly/blog/data-science-gender-race-disparity/

[25] Data available at https://www.ucop.edu/operating-budget/_files/ssgpdp/attachment_a_description_and_summary_table.pdf

Plan for Evaluating the Program

The program director will collect assessment data and evaluate two program learning outcomes (PLOs) each year. Institutional evaluation occurs as part of the regular program review cycle every seven years in accordance with UC Merced program review policy.[26] As the *Master of Data Science and Analytics* is a one-year program, assessing multiple PLOs each year will enable us to fully assess the program within the first three years, developing results, conclusions, and recommended actions to improve student learning as the program grows. An annual report will be circulated to the faculty and appropriate staff. Faculty will discuss the report at an annual meeting dedicated to that purpose and will prepare action plans to address any concerns or problems that are indicated by the assessment. Potential actions in response to findings might include revision to existing courses or their prerequisites, foci for student mentoring, and refined criteria for evaluating student learning. Annual learning outcomes reports will also be shared with the school dean and the Periodic Review Oversight Committee (PROC) to identify emerging trends in learning or assessment needs shared across programs for potential response at school or institutional levels.

Furthermore, the MDSA program will establish an Industrial Advisory Board (IAB) in collaboration with UC Merced's External Relations team that will be actively engaged in program assessment, student outreach, and developing potential employment opportunities.

The program will assess learning outcomes through multiple sources of evidence, including direct evaluations of student work in the context of program courses and indirectly through student and alumni surveys.

Program Learning Outcomes

The goal is for students to develop basic competence in the skills needed to understand and make use of data in a range of real-world problems, and for students to be ready to engage in lifelong learning and continuing education. Table 3 maps the PLOs to curriculum elements, and Appendix B describes the program rubrics. There are six PLOs:

1. **Methods.** Students will integrate appropriate quantitative, statistical, analytical, ethical, algorithmic and coding paradigms to identify knowledge management, planning and strategic decision-making solutions in different organizational and socio-technical situations. The program covers two core programming language (R and Python).
2. **Communication**. Students will selectively draw on different modes of communication (verbal, oral, graphical, code) to inform, engage and inspire in a clear and concise manner to diverse audiences comprised of experts and non-experts.
3. **Design.** Faced with a problem-oriented case, students will design resourceful and ethical informatics-based solutions by integrating data, methods and web-based technologies to strategically organize, manage, communicate, and deliver information-based services.
4. **Team.** Faced with a problem-oriented capstone project, students will gain experience integrating knowledge, skills, theory, and methods by leveraging team- and data-oriented

[26] https://assessment.ucmerced.edu/assessment-campus/annual-assessment/program-review

productivity solutions for sharing and integrating effort and data under time and other resource constraints.

5. **Ethics.** Students will understand the imperatives underlying research and data ethics, and will scope out the societal context and implications of their work by applying conceptual frameworks from the humanities and social sciences – to identify ethical, legal, and social issues surrounding data collection and analysis, to creatively develop and evaluate and implement responses to these issues; and to discover the embedded values and norms in data sets and algorithms, and to develop predictive skills for identifying emergent issues from the implementation or adoption of data science products or processes.
6. **Applications.** Students will apply informatics theory and data-management methods to address boundary-spanning problems pertaining to business, management, economics, sociology, psychology, cognitive science, environmental science and engineering.

TABLE 3: MDSA CURRICULUM MAP

Requirement/PLO*	Methods	Communication	Design	Team	Ethics	Applications
Required Courses						
Methods of Data Science I	I/M**	I	I	I	M	I/M
Data Ethics		I/M	I	M	M	M
Statistics and Probability	A		M		M	M
Methods of Data Science II	M/A	M	M	M	M	M
Visualization/Communication	M	A	A	I	M	A
Three Elective Courses	I/M	M/A	M/A		M/A	M/A
Capstone Activity	A/E	A	A/E	A/E	A	A

* All curriculum elements are required for successful completion of the program.

** **I** (Introduced) – The skills and knowledge associated with the PLO are introduced. Students meet, experience, begin to practice, and are assessed on the abilities through typical course activities.

M (Intermediate) – The skills and knowledge associated with the PLO are introduced. Students meet, experience, begin to practice, and are assessed on the abilities through typical course activities.

A (Advanced) - The skills and knowledge associated with the PLO are advanced. Through regular practice with feedback and related assessments, students continue to develop their proficiency.

E (Expert) - Through advanced assignments/assessments students further refine and then demonstrate the skills and knowledge associated with the PLO at a level of proficiency/competence expected of a graduate of the program. Though we will aspire to develop expert-level skills, we do not expect most students will attain this level in a one-year program.

ASSESSMENT PLAN

A basic elaboration of the assessment of the PLOs follows, which collectively enables the evaluation of the development of intermediate and advanced levels over the course of each student's completion of the MDSA. program. Table 3 maps the PLOs to curriculum elements. Appendix B describes the specific rubric for assessment of the PLOs.

1. **Methods.**
    - *Year assessed:* Program Year 1
    - *Direct evidence:* Written projects and oral presentations from Core courses, especially Data Science Methods I/II and Statistics & Probability; presentations, individual and team management reports, and short case reports from Capstone.

- *Indirect evidence:* Course grades from Data Science Methods I/II and Statistics & Probability; surveys of graduating students and alumni; group interviews of current students.
- *Performance targets:* For projects and presentations in coursework, it is expected that 100% of students will score "Intermediate" or better. It is expected that students will score "Advanced" or better on the presentations and projects for Statistics & Probability. Midway through the program, it is expected that students will rate themselves as "Intermediate" or better, and will rate themselves as "Advanced" or better at the end of the program.

2. **Communication**.
    - *Year assessed:* Program Year 2
    - *Direct evidence:* Written projects and oral presentations from Core courses, especially Visualization/Communication and Ethics; presentations, individual and team management reports, and short case reports from Capstone.
    - *Indirect evidence:* Course grades from Visualization/Communication and Ethics; surveys of graduating students and alumni; group interviews of current students.
    - *Performance targets:* For projects and presentations in coursework, it is expected that 100% of students will score "Intermediate" or better. It is expected that students will score "Advanced" or better on the presentations and projects for Visualization/Communication. Midway through the program, it is expected that students will rate themselves as "Intermediate" or better, and will rate themselves as "Advanced" or better at the end of the program.
3. **Design.**
    - *Year assessed:* Program Year 2
    - *Direct evidence:* Written projects and oral presentations from Core courses, especially Visualization/Communication and Capstone; presentations, individual and team management reports, and short case reports from Capstone.
    - *Indirect evidence:* Course grades from Visualization/Communication and Capstone; surveys of graduating students and alumni; group interviews of current students.
    - *Performance targets:* For projects and presentations in coursework, it is expected that 100% of students will score "Intermediate" or better. It is expected that students will score "Advanced" or better on the presentations and projects deliverables relating to their Capstone Activity. Midway through the program, it is expected that students will rate themselves as "Intermediate" or better, and will rate themselves as "Advanced" or better at the end of the program.
4. **Team.**
    - *Year assessed:* Program Year 3
    - *Direct evidence:* Written projects and oral presentations from Core courses, for instance, Data Ethics and Capstone; presentations, individual and team management reports, and short case reports from Capstone.
    - *Indirect evidence:* Course grades from Data Ethics and Capstone; surveys of graduating students and alumni; group interviews of current students.
    - *Performance targets:* For projects and presentations in coursework, it is expected that 100% of students will score "Intermediate" or better. It is expected that

students will score “Advanced” or better on the presentations and projects for Data Ethics and Capstone. Midway through the program, it is expected that students will rate themselves as “Intermediate” or better, and will rate themselves as “Advanced” or better at the end of the program.

5. **Ethics.**
    - *Year assessed:* Program Year 3
    - *Direct evidence:* Written projects and oral presentations from Core courses, for instance in Data Ethics and Data Methods I; presentations, individual and team management reports, and short case reports from Capstone. Hicks and Gunn plan to seek external funding to develop a robust, program-specific ethics assessment.
    - *Indirect evidence:* Course grades from Data Ethics and Data Methods I; surveys of graduating students and alumni; group interviews of current students.
    - *Performance targets:* For projects and presentations in coursework, it is expected that 100% of students will score “Intermediate” or better. It is expected that students will score “Advanced” or better on the presentations and projects for Data Ethics. Midway through the program, it is expected that students will rate themselves as “Intermediate” or better, and will rate themselves as “Advanced” or better at the end of the program.
6. **Applications.**
    - *Year assessed:* Program Year 1
    - *Direct evidence:* Written projects and oral presentations from Core courses, especially Visualization/Communication, Data Methods II and Capstone; presentations, individual and team management reports, and short case reports from Capstone.
    - *Indirect evidence:* Course grades from Visualization/Communication and Data Methods II ; surveys of graduating students and alumni; group interviews of current students.
    - *Performance targets:* For projects and presentations in coursework, it is expected that 100% of students will score “Intermediate” or better. It is expected that students will score “Advanced” or better on the presentations and projects for Visualization/Communication and Data Methods II . Midway through the program, it is expected that students will rate themselves as “Intermediate” or better, and will rate themselves as “Advanced” or better at the end of the program.

The Degree Program

Admissions Requirements and Preferences

The program is focused on serving recent graduates from any undergraduate major. Initial admissions will focus on recent graduates of UC Merced’s undergraduate programs, with a goal of generating a local data analytics workforce for the Central Valley and beyond.

Applicants admitted into the program will be expected to have completed a bachelor’s degree at a four-year accredited college or university and to have attained an undergraduate academic record that satisfies the standards established by the Graduate Division, University of California, Merced. All applicants whose first language is not English must also submit an acceptable TOEFL test score prior to admittance, scoring at least 550 on the paper exam or 80 on the

TOEFL iBT. A minimum undergraduate GPA of 3.0 is required; exceptions will be considered on a case-by-case basis. All applicants must meet the general requirements as set forth in the Graduate Studies section of the General Catalog. Applications will be accepted for admission in Fall semester only. GRE exam scores are **not** required.

In addition, a minor or major in a field of quantitative or mixed-methods empirical research (natural science, engineering, social or behavioral science, or humanities) *or* at least two years of work experience regularly using methods of quantitative analysis is preferred. Applicants lacking a strong quantitative or computing background will be strongly encouraged to participate in intensive pre-program "bootcamp" provided before the start of Fall semester.

Program of Study

The one-year Master of Data Science and Analytics will provide a unique educational experience to prepare students for increasingly common types of data problems encountered in real-world settings. Leveraging UC Merced's substantial depth in a range of areas related to data science and analytics – including visualization, remote sensing, systems modeling, data ethics, computational social science – the program will be unique, integrating foundational informatics practice with needed soft skills for data communication, open frameworks for team science, ethical use of data in visualizations and applications, and incorporating tracks in sustainability and environment, human behavior, and policy and decision making. We note that admitting students from a broad set of academic backgrounds presents challenges and also opportunities, intentionally stimulating an enriched and real-life learning experience. With a cohort of students coming from diverse academic majors and programs, challenges include finding effective ways to bring all students up to the same level of capability with background concepts and needed tools quickly. Yet at the same time, we have the opportunity to replicate closely in the classroom and in project groups problems of coordination and communication among diverse stakeholders, creating a learning environment that mirrors the real world. The objective is to have the cohort finish the program at the same level of competency and remain networked through their careers as an interdisciplinary and interactive alumni network.

The MDSA program requires a total of 33-36 semester units taken over one-year (including the summer capstone experience). These credits are distributed between a standard schedule of *core courses* (20 units), a choice of *theme* (9-12 units),[27] and a culminating *capstone* experience (4 units). *Core courses* focus on fundamental data science competencies, including exploratory data analysis, ethical analysis, and normative evaluation, statistical inference, and visualization and communication using both static and interactive media. Students select a *theme*, choosing three elective courses relevant to this theme from the set of approved courses, including courses in other graduate programs on campus, with one fulfilling the communication/visualization requirement. The selection of theme courses enables students to learn specialized analytical techniques (such as agent-based modeling or GIS), acquire domain knowledge and theoretical background relevant to their intended career path, and exercise data science skills in boundary-crossing application areas. The selection of theme courses gives the program its distinctive cross-disciplinary emphasis on real-world application of data science methods in *Sustainability and Environment*, *Human Behavior*, and *Policy and Decision Making*. The capstone activity is an integrative team-oriented project done during Summer. In addition, all students will have the

[27] Core courses are all 4 units, whereas theme courses may be 3 units or 4 units depending on program.

option of taking the Data Bootcamp offered shortly before Fall Semester. The Data Bootcamp will draw on open-source scientific computing curricular materials developed by The Carpentries[28] and the UC Merced Library, providing students who come in with less programming experience initial exposure to text interfaces and scripting. Table 4 illustrates the program of study.

TABLE 4: MASTER OF DATA SCIENCE AND ANALYTICS PROGRAM OF STUDY

Term	Title
Summer	Data Bootcamp (optional)
Fall	Core: Methods of Data Science I
	Core: Data Ethics
	Core: Statistics and Probability
	Theme: Elective 1
Spring	Core: Visualization/Communication
	Core: Methods of Data Science II
	Theme: Elective 2
	Theme: Elective 3
Summer	Capstone

All required courses must be taken for a letter grade, and students must receive a grade of B or better in all courses to complete the program successfully; less than a B is considered failing. Because the program is offered as a cohort model in one year, there will be no opportunity to repeat a failed course in the same year; students who fail one or two courses (except Capstone) will be given the opportunity to make these up the following year to complete degree requirements; students who fail Capstone will not be given the chance to make up the requirement; students who fail more than two courses will not be given the chance to make up the requirements. Any student may apply for readmission to the program.

REQUIREMENTS

Students must take courses that fulfill the following requirements: Methods of Data Science I and II, Data Ethics, Statistics, and Visualization/Communication. Courses that fulfill these requirements will be developed from current, recent, and near-future graduate program offerings in CIS, MIST, and ECON. Current core requirements may be fulfilled by existing (or soon to exist) graduate courses (see Table 5):[29]

[28] https://carpentries.org/

[29] Some of these are being offered in AY 2020-21 and AY 2021-22 as "topics" courses in their respective program; all will be proposed as standalone courses. Descriptions of the courses are provided below; sample syllabi are in Appendix D.

TABLE 5. CORE REQUIREMENTS AND COURSES

Requirement	Course
Methods of Data Science I	COGS 2XX: Data Science
Data Ethics	COGS 2XX: Data Ethics
Statistics and Probability	COGS 210: Statistics
Visualization/Communication	COGS 2XX: Global Good Lab
	MIST 232: GIS Decision Analysis
Methods of Data Science II	ECON 2XX: Data Science for the Social Sciences

In addition to core requirements, each student must choose one of three themes and complete three approved courses for that theme drawn from a variety of graduate groups on campus (see Table 6).[30]

TABLE 6. THEME COURSES[31]

Theme	Electives
Sustainability and Environment	COGS 222: Modeling Social Behavior
	ES/MIST 232 Applied Climatology
	ES 240: Water Resources Planning and Management
	ES 252: Remote Sensing of the Environment
	ES 260: Sustainable Energy
	ES 2XX: Life Cycle Assessment
	MIST 215: Political Ecology and Complexity
	MIST 254: Systems Thinking
	PH 204: Environmental Health
Human Behavior	COGS 204: Complex Adaptive Systems
	COGS 222: Modeling Social Behavior
	ECON 205: Applied Econometrics
	MIST 254: Systems Thinking
	MIST 261: Qualitative Methods for Management
	MIST 271: Network Science
	MIST 2XX: Service Innovation
	POLI 213: Experimental Methods in Political Science
	POLI 219: Behavioral Game Theory
	POLI 251: Political Cognition
Policy and Decision Making	COGS 222: Modeling Social Behavior
	ECON 205: Applied Econometrics
	MIST 254: Systems Thinking
	PH 202: Epidemiology
	PH 204: Environmental Health
	PH 216: Health Policy
	PSY 224: Health Disparities
	PSY 225: Health Risk Decision Making
	POLI 219: Model Based Inference
	POLI 224: Subnational Politics

[30] Other electives may be taken for credit by permission of instructor and graduate advisor.

[31] Theme courses were selected based on the following criteria: (1) offered regularly (annually or biannually), (2) not restricted to students in an existing graduate program, (3) covers analytical methods expected to be of interest to students the MDSA program. Other four-unit graduate courses may satisfy the theme requirements by petition to the program chair.

The capstone activity is an integrative team-oriented project conducted during the Summer session: Starting with an existing real-world problem identified with support of a project advisor (faculty, administration, industry), students will propose, design and develop a data-oriented solution; a viable product (e.g. interactive dashboard, entrepreneurial venture oriented around a data-service innovation, or data collection mobile app), report, and final presentation are required.

Normative Time to Degree

This is a one-year (twelve-month) program, including Fall and Spring semesters and a capstone project completed in the Summer semester. Students who do not successfully complete some requirements with their cohorts in the year of admission may choose to enroll in courses needed in the following year only; exceptions may be allowed on a case-by-case basis.

Licensing/Certification

None.

Projected Needs

Importance to the Discipline and Workforce

The fourth industrial revolution will transform our capacity to design and control cyber-physical systems that integrate breakthrough technologies in AI, virtual reality and internet-of-things.[32] In addition to the new products and services these technologies will provide, they will also contribute to deluge of information that is produced, recorded, and processed into valuable data.

With the *Digitization of Everything*,[33] the world has become primed for unabating data production and consumption. Yet the modern workforce has not yet developed the foundational skillset for fully operational on-demand data consumption. Furthermore, and in line with previous industrial revolutions, new conceptual and policy vacuums are emerging with respect to the justifiable and ethical use of data analytics, we require careful analysis of the nature and social impact of these new technologies. Technological advances that aid the process of producing data and extracting value are plenty – e.g. miniaturized and remote sensing for distributed collection; AI coupled with high-performance computing for scalable yet flexible solutions; and cloud computing and data management for distributed data analytics that increasingly builds on open-source computing tools, that together are transforming our ability integrate and analyze massive data sets.[34]

---

[32] National Academies of Sciences, Engineering, and Medicine, 2017. The Fourth Industrial Revolution: Proceedings of a Workshop—in Brief.

[33] Building America's skilled technical workforce. National Academies Press; National Academies of Sciences, Engineering, and Medicine, 2017. Information technology and the US Workforce: Where are we and where do we go from here?. National Academies Press; National Research Council, 2015. Promising practices for strengthening the regional STEM workforce development ecosystem. National Academies Press; National Academies of Sciences, Engineering, and Medicine, 2017. Strengthening Data Science Methods for Department of Defense Personnel and Readiness Missions. National Academies Press; National Research Council, 2013. Future US workforce for geospatial intelligence. National Academies Press.

[34] National Research Council, 2010. Steps toward large-scale data integration in the sciences: Summary of a workshop. National Academies Press; National Academies of Sciences, Engineering, and Medicine, 2020. Neuroscience Data in the Cloud: Opportunities and Challenges: Proceedings of a Workshop.

Consequently, despite highly scalable technological advances to facilitate both production and consumption, the workforce has lagged in absorbing the fundamental skillset for fully realizing *human-in-the-loop* cyber-physical systems. As such, the demand for informatics-literate workforce – comprised of teams of specialized analysts capable of identifying opportunities for data-driven problem solving, translating the problem-, solution- and decision-space to decision makers (ranging from organizational leadership to everyday citizens), and then designing and implementing feasible solutions – has never been in such high demand.

The proposed Data Science and Analytics program will integrate statistical, computational and visualization skills that are key dimension of this demand, in academia as well as both the private and public sectors. Upon completion, graduates of this one-year program will have complemented their deep undergraduate training with a broad and pragmatic informatics skillset, thereby increasing their number of potential career paths.

Upon graduation, some students may find it appealing to return to academia for an advanced research degree (M.S. or Ph.D.), as this multi-disciplinary program with focus on applied skills is likely to facilitate a sense of individual empowerment and a spawning of new interests. Such academic returnees would be well-equipped for either core or interdisciplinary studies. Based upon the integrative design of our proposed program, which features multiple tracks, students will be well prepared in contemporary data science paradigms calling for open data and code sharing protocols;[35] integration across massive cloud-based data sources;[2] and working within team environments on challenging socio-technical problems calling for interdisciplinary exploration, multidisciplinary integration and transdisciplinary vision.[36]

Other graduates will find employment in a range of private and public sectors that are developing and integrating large informatics pipelines within the organization — to aid with seamless real-time data collection and integration, risk management, accounting, assessment, auditing, and strategy, among other tasks. The proposed program will produce attractive employees equipped with foundational training and real-world problem-solving experience that will prove crucial in overcoming the paradoxical *skills-gap*[37] — whereby many employers expect experienced employees with advanced skillsets but are not willing to offer the foundational training to bridge the gap.

This program will facilitate bridging this gap by providing foundational skills delivered in 5 Core Courses that integrate across the program's six learning outcomes: *Methods*; *Communication*; *Design*; *Team*; *Ethics*; and *Applications*. For example, the core course *Methods of Data Science I* will introduce not only data management and analytics methods but will also teach students how to identify scenarios where a data-scientific approach is likely to outperform alternative assessment methods. This first semester course will be complemented by a course on *Data Ethics*, providing essential confrontation with the challenges and implications associated with collecting, securing, analyzing, and disseminating results based upon human-generated data. Indeed, to this end, a defining feature of the proposed program is that graduates will be grounded

[35] National Academies of Sciences, Engineering, and Medicine, 2018. Open science by design: Realizing a vision for 21st century research. National Academies Press.

[36] Committee on Facilitating Interdisciplinary Research, National Academy of Sciences, National Academy of Engineering and Institute of Medicine, 2004. Facilitating interdisciplinary research.

[37] Cappelli, P., 2012. Why good people can't get jobs: The skills gap and what companies can do about it. Wharton Digital Press.

in data chivalry, being forced to explore the full scope of potential outcomes that reflect upon environmental and humanist dimensions.

Second semester core courses in *Statistics*, *Data Communication and Visualization*, and *Methods of Data Science II* will train students in various empowering methods for extracting and communicating novel insights and value[38] from both structured (traditional tabular data) and unstructured (e.g. images) data sources using industry-standard platforms and technologies. Meanwhile, by providing three thematic tracks for topical specialization (in Sustainability and Environment, Human Behavior, and Policy and Decision-making) students will interact their skills with real-world scenarios, capitulated by the integrative Capstone project. This final project will be supervised by an external entity, e.g. possibly a UCM faculty or a local industry or non-profit partner, thereby reinforcing the human-human interaction dimension that is critical to problem-identification and decision-making that accounts for environmental sustainability and social justice impacts.

Graduates of the proposed MDSA program will be in high demand, in particular in non-technology sectors (including the public sector), which are projected to absorb 90% of IT-oriented jobs.[39] As such, graduates will be in close proximity to a variety of local industry hubs – from Los Angeles (media and entertainment, international trade) to San Diego (pharmaceuticals), San Francisco (Information Technology, high-performance computing, web and social networking services), and the Central Valley (agriculture).[40] This regional proximity is important, given that many first-generation students have additional family responsibilities and traditions that preclude studying outside of California.

Meeting the Needs of Society

The proposed program is designed around thematic tracks in *Sustainability and Environment*, *Human Behavior*, and *Policy and Decision Making* in an effort to address demand for skilled informatics workforce outside the traditional business and IT sectors. For this reason, the curriculum complements basic data science and analytics methods with essential *soft skills* tailored around effective data communication – methods of open science and team science; data dashboard development for interactive exploration and decision-making; and principles of data ethics. Such courses are not typically found in other M.S. in Data Science and Analytics programs.[41] Hence, graduates of the proposed program will develop data analytics intuition against the backdrop of environmental sustainability and social justice themes that are core to UC Merced's mission and vision. Likewise, students will be competitive in other emerging data-oriented job sectors such as Data Journalism that play an increasing role in democratizing data literacy across society.

---

[38] National Research Council, 2015. Training students to extract value from big data: Summary of a workshop. National Academies Press.

[39] Burning Glass Technology White Paper, 2019. Beyond Tech: The Rising Demand for IT Skills in Non-Tech Industries.

[40] National Academies of Sciences, Engineering, and Medicine, 2019. Improving Data Collection and Measurement of Complex Farms. National Academies Press.

[41] Other M.Sc. in Data Science and Analytics programs, typically offered by Computer Science and Statistics departments, tend to focus on training in sophisticated analytical methods and frameworks that require prior expertise in mathematical and statistical theory and object-oriented programming.

Also, because data analytics jobs can be performed remotely with relative ease and efficiency, graduates of the proposed program are likely to find employment in a resilient job market sector.[42] Such delocalization of the work office will help to address the financial and health burdens associated with living in megacities, such as affordable housing shortages and contagion control. As of 2021, roughly 36% of wage and salary workers worked at home on average 5.71 hours per week,[43] which was particularly high for employees from the information industry,[44] representing a growing trend[45] over the last decade as functional teleconferencing technology and cloud computing infrastructure[46] become widely available; moreover, recent psychological research indicates that remote work can have positive impacts on `productivity, creativity and morale'.[47]

And finally, as a designated Hispanic (HSI), Minority (MSI) and Asian American and Native American Pacific Islander (AANAPISI) -serving institution, with 71% of undergraduates reporting as first-generation students,[48] UC Merced is the most diverse UC campus.[49] As such, the proposed program will contribute to broader campus, state and nationwide efforts to develop STEM pathways that provide diversity-oriented value-added to the workforce,[50] in particular relating to informatics and computing.[51]

Relationship to Research and Professional Interests of the Faculty

Many faculty members in the core graduate groups of the proposed program – CIS, ECON, and MIST – have research interests in the broad area of data science and analytics. These faculty represent a substantial pool of academic Capstone supervisors who can readily identify a problem that calls for data-oriented solution. For example, in collaboration with computer scientists, ECON faculty member A. Johnston co-developed *PrivateJobMatch*, a data privacy-oriented algorithm for improving job-market recommender-system match quality. [52] MIST faculty member J. Abatzoglou co-developed *Climate Engine* which integrates cloud computing

---

[42] Data from Glassdoor, Inc. cited in: Say goodbye to six-figure starting salaries – with these exceptions, CNBC. (6/8/2020)

[43] US Bureau of Labor Statistics: https://www.bls.gov/news.release/atus.t07.htm

[44] US Bureau of Labor Statistics: https://www.bls.gov/news.release/flex2.t01.htm

[45] Global Workplace Analytics estimates that roughly half of US jobs are at least partially compatible with remote work, and that 25% to 30% of US employees will continue to work from home once a week or more after the COVID-19 pandemic.

[46] National Research Council, 2010. Steps toward large-scale data integration in the sciences: Summary of a workshop. National Academies Press.

[47] The future of remote work. Zara Abrams, American Psychological Association (2019)

[48] https://visualizedata.ucop.edu/t/UCMerced/views/EnrollmentStatistics/EnrollmentFacts/fd169f6c-4fd6-435d-a473-bdabebf4b990/4626c818-c7ad-4363-b179-76fc3c50a25e?:display_count=n&:showVizHome=n&:origin=viz_share_link

[49] https://www.ucmerced.edu/diversity

[50] National Academies of Sciences, Engineering, and Medicine, 2019. Minority serving institutions: America's underutilized resource for strengthening the STEM workforce. National Academies Press.

[51] Building America's skilled technical workforce. National Academies Press; National Academies of Sciences, Engineering, and Medicine, 2017. Information technology and the US Workforce: Where are we and where do we go from here?. National Academies Press; National Research Council, 2015. Promising practices for strengthening the regional STEM workforce development ecosystem. National Academies Press; National Academies of Sciences, Engineering, and Medicine, 2017. Strengthening Data Science Methods for Department of Defense Personnel and Readiness Missions. National Academies Press; National Research Council, 2013. Future US workforce for geospatial intelligence. National Academies Press.

[52] Github: PrivateJobMatch – a Python tool for simulating a Job Market; Saini, A., Rusu, F., & Johnston, A. (2019). PrivateJobMatch: a privacy-oriented deferred multi-match recommender system for stable employment. In *Proceedings of the 13th ACM Conference on Recommender Systems*.

solutions to integrate and visualize climate data using Google Earth Engine.[53] And CIS faculty member L. Padilla leads the *Spatial Perception, Applied Cognition & Education Lab*,[54] which conducts research at the interface of cognitive and visualization science. Broadly speaking, the majority of CIS, ECON and MIST faculty apply computational methods to empirical data to evaluate policy and model human-oriented phenomena emerging in socio-technical systems. Together, these complementary methods and data-oriented contexts give rise to a wide array of real-world informatics perspectives that are a strength of the proposed program.

Program Differentiation

As described previously, there are many data science and business analytics programs in the UC system (see Table 2), including programs run through engineering schools that focus on data science, programs run through management schools that focus on business analytics, and online programs that focus on data science. As suggested, our proposed program would be unique inside and outside the UC System as it straddles the line between deep data science education and applied analytics training, with a wide choice of thematic concentration for students. No other program we have found provides students the opportunity for broad interdisciplinary data science and analytics education proposed here. As data science is a highly interdisciplinary field, the configuration of this program at the nexus of three fields is a distinguishing feature compared to similar programs offered at other UC campuses.

External and Competitive Market Analysis

We engaged UC San Diego Extension Center for Research and Evaluation (CR+E) to provide an independent assessment of the competitive landscape, employment prospects, and market demand for our proposed Master of Data Science and Analytics (see Appendix E). In short, the assessment is very positive about the prospects for our proposed program. Though there are many master's programs in data science and analytics in California and across the country, the demand for data scientists is continuing to grow rapidly, with the number of job postings requiring data science skills increasing more than a 600% increase since 2010 and with the number jobs that require data science skills seven times greater than the number of jobs with data science titles. California has greater demand for data science skills than any other state. Surveying hundreds of college graduates, including a large number of UC Merced alumni, CR+E found almost all were interested in the MDSA program, with 85% somewhat likely or very likely to apply, with most of the recent college graduates expressing interest. Critically, there was broad interest across the three themes – and by graduates with backgrounds in engineering, social sciences, and many other areas. Survey responses indicated that at $45K, the proposed MDSA program would be reasonably priced for most.

Based on the CR+E Market Analysis, we selected $45K as a price-point that balanced the respondents' beliefs about program costs. For example, Figure 1 shows the proportion of 301 nationwide respondents (39% from California) who indicated their beliefs about the proposed

---

[53] ClimateEngine.org – a Google Earth Engine tool for on-demand processing of satellite and climate data via a web browser; Huntington, J., Hegewisch, K., Daudert, B., Morton, C., Abatzoglou, J., McEvoy, D., and T., Erickson. (2017). Climate Engine: Cloud Computing of Climate and Remote Sensing Data for Advanced Natural Resource Monitoring and Process Understanding. Bulletin of the American Meteorological Society

[54] *Spatial Perception, Applied Cognition & Education Lab* at UC Merced

tuition being too cheap, a bargain, premium, or too expensive. The majority (~58%) reported that $45K would be a bargain (shown in grey). Roughly 12% indicated that any price over $45K was "Expensive but still would consider," and 8% of the respondents indicated that over $45K was "Too Expensive to consider." Responses assessing the quality, as opposed to the affordability, also indicated support for $45,000.[55]

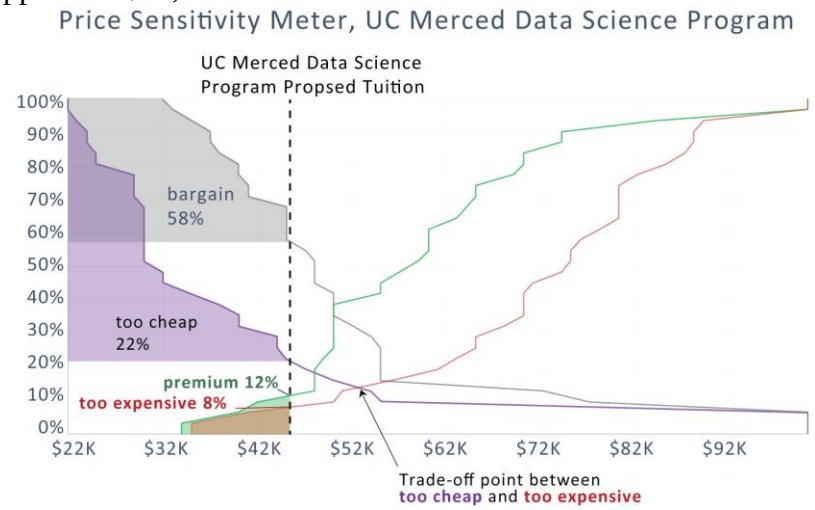
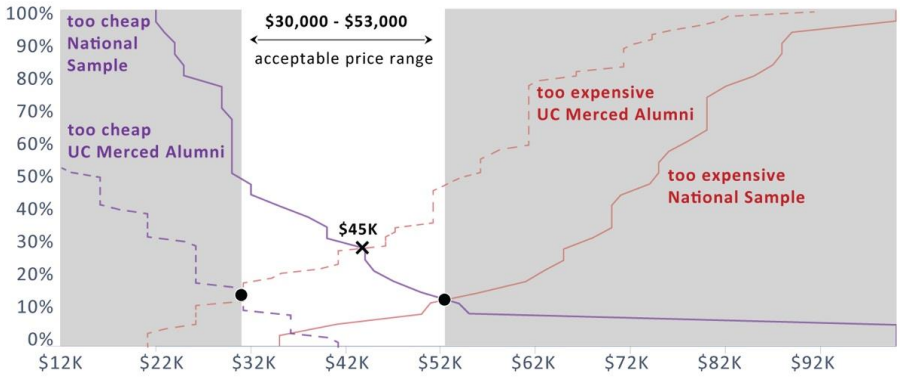


Figure 1. Annotated version of a figure from the UC San Diego Extension Center for Research and Evaluation Market Analysis. Figure shows the proportion of US respondents who indicated their beliefs about the proposed tuition being too cheap, a bargain, premium, or too expensive.

We selected the price point of $45K (rather than the trade-off point between too cheap and expensive of $52K, as shown in Figure 1) because we sought to balance the perspective from the national sample with the UC Merced Alumni responses. As shown in Figure 2, when comparing beliefs about the program tuition being too cheap and too expensive between the national sample and UC Merced alumni, there is a range from $32K to $52K that balances beliefs from both groups.



[55] Conversely, only 22.6% (respectively ~61%) of the respondents considered any price less than 45k to be "*Too cheap to be of good quality" (resp., "A Bargain or great buy"),* indicating that the $45,000 target cost is balanced around the bargain and premium sides.

Figure 2. Acceptable price range using the national and UC Merced alumni responses based on the Van Westendorp price sensitivity method.[56] Data from the UC San Diego Extension Center for Research and Evaluation Market Analysis.

Note that a price point of $42K would be closer to an exact compromise between the beliefs about the cost of both groups. Rather than reducing the tuition as a whole by $3K, we created a funding model that prioritizes lowering the cost to low-income students. In this plan, 20% will be returned to aid, which is a far larger percentage than other programs we have found (the next largest being 3% return-to-aid). If we assume that half the enrolled students have need for financial aid, then 20% return-to-aid would provide about $20,000 per student with need. By prioritizing need-based support, rather than reducing the overall tuition, the MDSA program will be affordable for UC Merced Alumni and other low-income students while maintaining a perceived high quality by the greater California and US populations.

Student Demand

The CR+E report assessed demand through a survey of college graduates, including UC Merced alumni (see Appendix E). Over 400 graduates were surveyed, most in California. More than 80% expressed interest in the program, with most suggesting that data science knowledge and skills gained would enhance job security and promote career advancement. The program was particularly appealing to recent social science and engineering graduates. There was broad interest expressed across all three theme areas. Overall, more expressed interest in a hybrid program than in either an online of an in-person program; we expect our core courses will adapt to a hybrid format over time, combining in-person and online components.

Opportunities for Placement of Graduates

A recent IBM report[57] has identified data science and analytics as one of the most rapidly growing career fields. This report characterizes a wide range of applications and domains for data scientists, including in finance, professional services, IT and other areas. And whereas about 40% of the jobs would require an advanced degree, such as a master's in data science and analytics, most of the jobs could be filled by graduates with a bachelor's degree in data science and analytics. Many rapidly growing, high-paying opportunities are available to data science graduates, including in business intelligence, information architecture, market research, and much more.[58] The Bureau of Labor Statistics (BLS) estimates the growth in business occupations to be about 10% and management occupations to be 8.5% from 2016 to 2026.[59] BLS estimates the growth in computer and mathematical occupations to the 14% in the same period, showing strong growth in the areas related to our program. According to an IBM report examining the demand for data scientists, 39% of job postings for data scientists or advanced analysts require an advanced degree.[60]

---

[56] https://www.forbes.com/sites/rebeccasadwick/2020/06/22/how-to-price-products/?sh=10a801855c75

[57] See https://www.ibm.com/downloads/cas/3RL3VXGA

[58] See, for example, https://www.onetonline.org/link/summary/15-1199.08, https://www.onetonline.org/link/summary/15-1199.02, https://www.onetonline.org/link/summary/13-1161.00, and https://www.onetonline.org/link/summary/15-1111.00

[59] Bureau of Labor Statistics, 2017. Employment Projections: 2016-2026. https://www.bls.gov/emp/ep_table_101.htm

[60] IBM, Burning Glass Technologies, and Business-Higher Education Forum, 2017. "The Quant Crunch. How the demand for data science skills is disrupting the market." https://www.ibm.com/analytics/us/en/technology/data-science/quant-crunch.html

The CR+E report similarly characterized the employment outlook for data science graduates very positively (see Appendix E). The clear need for data science skills across a variety of occupations aligns with the MDSA program's focus on skills, team projects, and applications. California has had the greatest number of job postings requiring data science skills.

Given these projections, it is expected that the graduates of the proposed Master of Data Science and Analytics program will find jobs that will make use of their newly earned knowledge and practical skills. During the first two years, the program will hire a student service staff member (50% FTE) to help the students find internship and job opportunities. As the MDSA enrollment increases over the years, there will be additional staff dedicated toward finding employment opportunities for students.

Faculty

The three primary graduate groups of the program – CIS, ECON, and MIST – currently include 46 members. Here we have included short research descriptions for each core faculty member of these groups (links to CVs are in Appendix J). Obviously, not all faculty across the three groups have deep interest in data science and analytics or are expected to teach in the program. Those we expect to be most involved in the MDSA program are marked with an asterisk (*).

Cognitive and Information Sciences

Zenaida Aguirre-Muñoz, Professor. Zenaida Aguirre-Muñoz's research integrates cognitive science, language, learning sciences, and assessment applied to the following research interests: (a) STEM education; (b) model-based assessment and instruction of dual language learners; (c) the impact of opportunity to learn on learning and achievement; and (d) content-area literacy development for dual language learners. Her research projects have been funded by organizations such as NSF, NIH, Department of Energy, and the US Department of Education.

Kristina Backer, Assistant Professor. Kristina Backer's research lies at the intersection of cognitive and auditory neuroscience. Real-world listening environments are often noisy. In these acoustically-adverse listening situations, perceptual and cognitive processes dynamically interact, thereby enabling successful speech comprehension. Thus, one line of research aims to elucidate the neural mechanisms underlying these dynamic perceptual-cognitive interactions in young adults with normal hearing. Age-related changes in auditory perception and certain cognitive processes, such as selective attention, can exacerbate the effort required for speech comprehension in noise. Therefore, a second research aim is to understand how aging and hearing loss affect the perceptual and cognitive processes that are engaged during listening. Finally, a third research goal is to characterize how language experience and language context modulate perceptual and cognitive processing during listening, throughout the lifespan. Methods used in the laboratory include behavioral paradigms and electroencephalography (EEG) recordings.

Ramesh Balasubramaniam, Professor.* Ramesh Balasubramaniam's research focuses on cognitive, neurophysiological and dynamical systems approaches to coordinated human action, using the human motor system as a vehicle to understand the embodied nature of cognition. In particular, Balasubramaniam's work looks at several functions of the motor system including planning, prediction, anticipation, and simulation to understand sensory and cognitive functions

of the brain. Methods include 3-D motion capture, eye tracking, robotic exoskeletons, TMS, EEG and other electrophysiological recordings to study the brain and behavior. Particular problems relate to (1) sensorimotor timing (2) visual and auditory guidance of action (3) posture, balance and the control of unstable objects (4) music and synchronization and (5) human interaction with robotic systems and agents. Much of this research borrows from principles of complex systems, where the behavior of a system of many components is significantly different, yet simpler than the behavior of the components themselves. This research is supported by several grants from the National Science Foundation and other sources. Balasubramaniam also directs the NSF-NRT graduate interdisciplinary training program in intelligent adaptive systems, which provides students with a learning environment that combines elements of computation and data science.

Heather Bortfeld, Professor. Heather Bortfeld's research follows two converging lines, one in typical and one in atypical language learning and use, specifically under adverse listening conditions. The first line looks at things like: how typically developing infants come to recognize words in fluent speech and the extent to which the perceptual abilities underlying this learning process are specific to language. The second line looks at the influence of perceptual, cognitive, and social factors on language development in deaf children who learn language through a cochlear implant. The acoustic signal these children hear through their implant is substantially degraded relative to the original source. Some do better using this signal than others. Bortfeld examines possible sources of variability in this learning process and how language outcomes can be maximized for this population. She also studies how postlingually deafened adults who opt for a cochlear implant learn to process the input. Bortfeld's approach is best characterized as integrating (1) multiple methods, (2) different levels of analysis, and (3) a broad theoretical perspective. The findings from this research underscore the central role of experience, both perceptual and social, in language learning and language processing throughout the lifespan.

Hanna Gunn, Assistant Professor.* Hanna Gunn is an Asst professor in the Department of Cognitive and Information Sciences at UC Merced. Gunn's research falls broadly within the area of "social philosophy" (social epistemology and feminist philosophy). Gunn's current research projects span three main topics: a theory of epistemic community and epistemic respect; social epistemic research into phenomena including filter bubbles, echo chambers, and polarisation; and understanding responsibility for manipulation and other violations of autonomy online.

Daniel Hicks, Assistant Professor.* Daniel Hicks is a philosopher of science, data scientist, and science policy researcher. Hicks' primary academic research focuses on the role of ethical and political values in science and public scientific controversies. Hicks is also interested in bibliometric, the use of statistics in scientific practice, and data ethics. As a data scientist, much of Hicks' work involves network analysis and multilevel generalized linear regression, including Bayesian models in Stan. Hicks has experience with topic models and word embedding's for text data, as well as spatial Durbin models and other methods for spatial data. Hicks works primarily in R, but also has experience with Python.

Colin Holbrook, Associate Professor.* Colin Holbrook is an Associate Professor of Cognitive and Information Sciences at the University of California, Merced. His program of research explores decision-making under contexts of threat, with particular focus on aggression, coalitional psychology, morality, and the representation of mental states. Holbrook's doctoral work explored processes by which threat cues heighten reactivity to emotional stimuli, particularly when individuals are unaware of having entered a state of threat-vigilance,

precipitating a number of judgment biases including heightened group prejudice. At the level of proximate neurobiological mechanisms relevant to such context-sensitive shifts, Holbrook employs transcranial magnetic stimulation to explore the role of the posterior medial frontal cortex region in modulating empathic and prosocial responses contingent on perceived group membership. In complementary research, Holbrook's research has examined the psychological mechanisms by which individual differences in threat-sensitivity related to political orientation or religious belief predict group prejudice and tendencies to aggress. Most recently, Holbrook has extended his work on group bias to human-robot interaction with anthropomorphic agents, systematically manipulating characteristics such as the apparent gender and ethnicity of a robot to assess potential effects on conformity with the robot's recommendations, and perceptions of the robot as possessing mental attributes such as emotion and intelligence.

Carolyn Dicey Jennings, Associate Professor. Carolyn Dicey Jennings' primary research project uses findings from cognitive science to explore the nature of attention and its impact on the mind. Distinctive about Jennings research is the claim that attention provides evidence of a self, which is responsible for directing attention, and that attention is a key element of perception, consciousness, action, and responsibility. Jennings also has a secondary, data-based research project on employment outcomes for doctoral students. An important aspect of this work is the focus on diversity and inclusivity and how it intersects with the graduate experience.

David Jennings, Associate Teaching Professor. David Jennings works primarily in ancient moral philosophy. Jennings research concerns how ancient philosophers understood the relationship between the individual and common good. Jennings is particularly interested in Plato and Aristotle's claim that philosophizing makes one not only a better human being but also a better friend, family member, and citizen. Similar questions can be found in contemporary practical ethics, and Jennings has an ongoing project concerning how much our obligations to those in need ought to limit our pursuit of other values, such as knowledge and beauty.

Chris Kello, Professor.* Professor Kello studies how people coordinate their speech and movements to work together in pairs and groups. His lab and collaborators use analysis methods from complex systems research to study how human behavior is timed and structured, and how this temporal structure reflects underlying thought processes. The types of coordination he studies include verbal and non-verbal interactions, and individual and group search processes. Kello also works on energy-efficient methods for machine learning that may be applied to edge computing applications such as smart remote sensing and signal analysis.

Tyler Marghetis, Assistant Professor. Tyler Marghetis's research explores how both the diversity and the discipline of abstract thought emerge naturally from the nested complex systems in which cognition occurs— the 'ecologies of cognition' — from solitary brains, to interacting bodies, to sociotechnical systems. Drawing on data both 'big' (large corpora of cognitive activity) and 'bespoke' (evidence from cross-cultural fieldwork), he focuses on cognitive domains that often settle into person- or community-specific ways of talking and thinking, before shifting abruptly to new regimes (e.g., mathematical reasoning, musical improvisation). This moves us closer to understanding, in general, when and why our abstract thought settles into stable regimes — and sometimes undergoes radical revolutions, from sudden religious conversion or mathematical insight, to language change over historical time.

Teenie Matlock, Professor, McClatchy Chair in Communications. Teenie Matlock is the McClatchy Chair in Communications, and a Professor of Cognitive Science. Her research

focuses on communication. Much of her work focuses on non-literal language, for instance, metaphors used in the expression of time, numbers and math, social distance, spatial configurations, and internet use. Matlock's recent work on metaphor examines its use in public discourse, for instance, in discourse about political campaigns, wildfires, climate change, and disease. Some of her work examines how people interpret grammatical information, for instance, tense and aspectual markers as well as evidentials. Other research has focused on how people produce and interpret manual gestures, for instance, in describing quantity. Some of Matlock's early work examined communication in the domain of human-centered computing.

David Noelle, Associate Professor.* David Noelle's research largely involves the fabrication, analysis, and testing of computational models of the cognitive processes and neural mechanisms that give rise to controlled behavior and explicit learning. Recent work has focused on the role of prefrontal cortex, and Associated brain areas, in working memory, cognitive control, and category learning. These models have been used to explain cognitive flexibility, the behavioral effects of frontal damage, aspects of instructed learning, and behavioral differences in people with autism spectrum disorders. In addition to his computational neuroscience work, Noelle conducts laboratory experiments on human learning and research on brain-inspired machine learning methods.

Lace Padilla, Assistant Professor.* Lace Padilla's empirical research examines the underlying cognitive mechanisms used in decision-making with visualizations, particularly during hazard events such as hurricanes and flash floods. This research focuses on evaluating the impact of basic cognitive processes (e.g., perception, attention, working memory, and knowledge) to create visualizations of data that are easier for people to use during high-risk events with uncertainty (e.g., evacuating before a hurricane strike). She works collaboratively with computer scientists, climate scientists, and anthropologists. Additionally, she is a Disaster Risk Management and Behavioral Science Consultant for the World Bank, where she applies her empirical research to support disaster risk managers in countries such as Haiti and Bangladesh.

Rachel Ryskin, Assistant Professor.* Rachel Ryskin studies how individuals achieve efficient language processing in the face of ambiguity, variability, and noise. She uses eye-tracking and EEG to examine how people use various sources of information (visuo-spatial perspective, theory of mind, language statistics, etc.) to constrain the real-time interpretation of spoken language, as well as the learning and memory processes that underpin these representations.

Paul Smaldino, Associate Professor.* Paul Smaldino is Assoc Professor of Cognitive and Information Sciences. His work employs mathematical and computational modeling to answer questions about social behavior and cultural evolution of humans and other animals. He is also known for his work modeling the population dynamics of scientific communities.

Michael Spivey, Professor.* Michael J. Spivey earned his B.A. in Psychology at UC Santa Cruz, and then his Ph.D. in Brain and Cognitive Sciences at the University of Rochester. After being a professor of psychology at Cornell University for 12 years, he came back to the UC system to help build the Cognitive & Information Sciences Department at UC Merced. By recording eye movements and reaching movements in natural tasks, and modeling the results with neural network simulations, his work shows that visual perception and language comprehension continuously interact with each other, and also with motor movement and situational context. Rather than treating cognition as separate computational modules devoted to vision, or language, or action, the dynamical system approach suggests that the mind is a richly

interactive process, inseparable from the activity of the body in its environment. Spivey's research program has published its findings in Science, Current Biology, PNAS, Trends in Cognitive Sciences, and many other top-tier journals. In 2010, Spivey received the William Procter Prize for Scientific Achievement from the Sigma Xi Society for Research. This research is described in his 2007 book, The Continuity of Mind (Oxford U. Press), and in his 2020 book, Who You Are (MIT Press).

Jeff Yoshimi, Associate Professor.* Jeff Yoshimi specializes in philosophy of mind and cognitive science, phenomenology (especially Husserl), neural networks, dynamical systems theory, and visualization of complex processes.

ECONOMICS

Catalina Amuedo-Dorantes, Professor. Catalina Amuedo-Dorantes research areas include labor economics, international migration and remittances. She has published on contingent work contracts, the informal work sector, international remittances, as well as on immigrant savings, health care and labor market outcomes. Her current research broadly focuses on immigration policy and its consequences. She examines the impact that state and local level immigration policy is having on the employment, education, fertility and human rights of undocumented immigrants, as well as on the effect of immigration policy geared towards high-skilled immigrants.

Briana Ballis, Assistant Professor. Brianna Ballis works on topics related to education and inequality. Her job market paper explored the long-run impacts of special education programs. She has also studied the impact of DACA on educational spillovers among peers. Her current work in progress explores the role of infant health on educational outcomes.

Christian Fons-Rosen, Associate Professor.* Christian Fons-Rosen's research mainly focuses on two branches: (1) economics of science; (2) firms and innovation. In (1), Fons-Rosen uses large datasets and statistical techniques including text analysis to track the evolution of science over time, to evaluate the effort of lower communication costs on scientific progress, and on the relative merits of doing science in universities versus corporations. Future work includes evaluating the impact of Mccarthyism on the direction of science and also analyzing the effort of an aging scientific cohort on the renovation of ideas. In (2), Fons-Rosen looks at the intensity of knowledge flows across firms and the role of geographical and institutional boundaries in the intensity of idea dissemination. This again implies using large datasets on the population of firms in many European countries matched to all their patent stock and, for example, tracking inventors over time and space.

Rowena Gray, Associate Professor.* Rowena Gray is an economic historian specializing in nineteenth and twentieth century United States. She investigates the effect of technological change on labor markets during the second Industrial Revolution which introduced electricity and related general purpose technologies. She also considers the first global era, assessing the impact of immigrants on crime rates, the housing market and innovation. She has published in top economic history outlets and is currently PI on a Russell Sage Foundation grant and co-PI on a NSF grant which will create new measures of housing prices and living standards over the long run in the urban United States. She enjoys incorporating undergraduates in her research and promoting her results in outlets such as the Conversation.

Zack Grossman, Associate Professor. Zack Grossman is a microeconomist who studies the impact of social and psychological motivations, as well as cognitive phenomena, on economic decisions. His approach is behavioral, reflecting an openness to how insights from psychology can help us understand economic behavior using both experimental and theoretical methods.

Justin Hicks, Associate Teaching Professor. Justin Hicks is an applied macroeconomist working with patent and publication data to identify the unintended spillovers of collaboration across international borders. Hicks loves teaching and being in the classroom and looks forward to continuing his own learning experience both in the classroom and in his research as he develops his career at UC Merced.

Rob Innes, Professor, County Bank Chair in Economics. Rob Innes conducts research in experimental economics with a focus on moral preferences; environmental economics with a focus on efficient regulation of pollution; theories of market structure, particularly in food markets and how different types of market competition affect prices and performance; and the law and economics of property rights, liability, and nuisance regulation.

Andrew Johnston, Asst Professor.* Andrew C. Johnston is an economist who specializes in public economics and labor. He studies the influence of social insurance and taxation on firm and worker behavior using natural experiments. For example, he studied the effect of extended unemployment benefits on search and job finding among unemployed workers. Motivated to identify policies to promote economic mobility, he has begun a research agenda in human-capital formation. Since teachers are the central input in public schooling, he has focused on studying the labor market for teachers, including the causes of teacher shortages, how to attract and retain excellent teachers, the role of compensating differentials in promoting equal opportunity, and how compensation and working conditions shape the quality distribution of teachers. His work has been published in the Journal of Political Economy and the American Economic Review: P&P. In 2016, He earned his PhD in economics from the University of Pennsylvania and his wife and children are his greatest joy.

Jason Lee, Associate Teaching Professor. Jason Lee's research has focused on financial markets and economic history.

Jesus Sandoval-Hernandez, Associate Teaching Professor. Jesus Sandoval-Hernandez is an Asst Teaching Professor (LPSOE) of Economics at the University of California Merced. He earned a PhD in International Economics from the University of California, Santa Cruz. Sandoval-Hernandez's teaching experience include Econometrics, Corporate Finance, Introduction to Business and Finance, The Economics of Emerging Markets, Statistical Inference, International Macroeconomics/Finance, Introductory Finance, Economic Development, International Microeconomics, Intermediate Macroeconomics, Latin American Economies, The Economics of Globalization, Economics and Management of Technology and Innovation, Personal Finance, and Introductory courses in Micro and Macroeconomics.

Kurt Schnier, Professor, Interim Vice Chancellor and Chief Financial Officer. Kurt Schnier studies policy analysis in the fields of health and resource economics. His research includes a focus on the response of physicians and hospitals to current and impending regulations that impact the provision of health care as well a focus on facilitating the development of efficient marine resource policy.

Theofanis Tsoulouhas, Professor. Fanis Tsoulouhas' research focuses on applied theory, particularly contract theory and the economics of information. Most of his research deals with the implications of asymmetric information and moral hazard problems (which arise when agents undertake unobservable actions with stochastic effects). His earlier work considered, first, the limitations imposed by limited liability, bankruptcy and commitment problems on financial and labor contracts in order to show the optimality of debt as opposed to equity contracts and, second, the strategic role of information gathering and transmission. His later work has focused on organizational design and on applications of absolute or relative performance evaluation (i.e., tournaments and contests) in areas such as executive promotion and contracts under moral hazard, limited liability for the principal or liquidity constraints for the agents and/or heterogeneity of agents. His latest published papers deal with performance pay incentives and offshoring, and with the optimality of tournaments under adverse selection (which arise when some parties have private information when contracts are signed). His latest working paper examines why countries in financial distress strategically delay seeking help. He is also working on entrepreneurship, effort and top incomes in a small open economy, and on business commonality, standardization and offshoring.

Ana Tur-Prats, Asst Professor. Ana Tur-Prats is an applied economist, with broad research interests in Culture, Gender and Economic History. Tur-Prats' research mainly focuses in two branches: (1) violence against women; (2) social capital. In (1), Tur-Prats has analyzed the long-term determinants of intimate-partner violence by focusing on the historical family structure that was prevalent in the past. She has also looked at how individuals' responses to changes in the gender gap in unemployment in terms of intimate-partner violence are shaped by their underlying cultural norms about the appropriate role of men and women in society. In current work Tur-Prats is analyzing how cultural norms can explain the prevalence and intensity of conflict-related sexual violence. In (2), Tur-Prats looks at the historical determinants and long-term persistence of social capital, by focusing on the legacy of the commons; and also analyzes how social capital has a positive impact on local economic development today. On a separate project, Tur-Prats looks at how conflict can actually destroy social capital, by examining the long-term consequences of the Spanish Civil War on generalized trust and other outcomes.

Greg Wright, Associate Professor. Greg Wright has a PhD from the University of California, Davis and studies international trade, immigration and technological change, primarily focusing on the labor market impacts of each. His current projects explore the long-run consequences of international trade shocks and the role of patents in mediating global trade flows. His work has appeared in top Economics journals such as the American Economic Review, the American Economic Journal: Applied Economics, the American Economic Journal: Economic Policy, and the Journal of International Economics.

MANAGEMENT OF INNOVATION, SUSTAINABILITY AND TECHNOLOGY

John Abatzoglou, Professor.* John Abatzoglou studies climate science and climate impacts in the American West. His group additionally develops monitoring and forecasting technologies to improve the climate readiness of human and natural systems.

Roger Bales, Professor. Roger Bales focuses on the scientific research of mountain hydrology and biogeochemistry, polar snow and ice, climate impacts and water resources. He applies his

research to real world problems and engages with decision makers to improve information and investments in California's water resources systems.

Anita Bhappu, Associate Professor. Anita Bhappu has published research on team collaboration, workplace diversity, service delivery and digital technology. She is currently studying platforms in the sharing economy. Her research examines technology affordances and the nature of service interactions on ridesharing and shared lodging platforms, particular from the provider perspective. She is also investigating employee trust and engagement related to coworker collaborative consumption in a field study of an organization-sponsored sharing platform.

Spencer Castro, Assistant Professor.* Spencer Castro studies the performance limitations of human-machine systems. He focuses on the validity of behavioral and physiological metrics as measures for different aspects of workload, as well as for quantifying the risk of adverse outcomes due to these workload metrics in driving.

Jeff Jenkins, Assistant Professor.* Jeff Jenkins' research aims to answer questions such as, What is the visitor carrying capacity of our national parks? How are social equity and ecological conditions valued and negotiated through public process? What features contribute to visual buy-in for climate adaptation? How do political-legal, economic, aesthetic, and historic land use considerations shape multiple use landscapes? Are the current and future trajectories of forest management bound by original mandates or can new paradigms emerge amid socio-ecological complexity? Jenkins works with partners at the National Park Service, U.S. Forest Service, and other public lands to inform management decisions through research, teaching, and public facing data collection methods. Technologies including geovisualization and geographic information systems allow for information to be geospatially communicated in much of this work. This research is built upon theoretical traditions of critical resource geography, socio-ecological systems, political ecology, regional planning, and public lands management.

Catherine Keske, Professor. Catherine Keske is an agricultural and resource economist who conducts interdisciplinary coupled natural human systems research at the food-water-energy nexus. Keske's projects often involve disadvantaged communities in mountainous, semi-arctic, or arid ecosystems seeking economically viable, sustainable environmental management. Several of my popular studies involve food security and food sovereignty; sense of place; conservation easements; the economic value of Colorado Fourteeners; and, food, energy, and biochar co-production. Keske publishes in a variety of literatures, including economics, engineering, sociology, law, public planning, and sustainability science. Keske uses a continuum of quantitative and qualitative research methods and enjoy working on international projects. Presently, Keske is working on two multi-investigator projects funded by California's Strategic Growth Council to mitigate climate change. One project aims to reduce methane emissions by developing an economic viable system for biochar production and manure management at Central Valley dairies. A second project aims to improve ecosystem service valuation of California's forests for their carbon storage and water conservation benefits. Keske is also leading a project to establish an Engineering Research Center at UC-Merced for precision agriculture and agricultural technology, and another project to develop a sustainability index that will promulgate sustainable dairy management practices in the semi-arid U.S. West.

Crystal Kolden, Associate Professor.* Crystal Kolden is a pyrogeographer who studies wildfire interactions with socio-ecological systems broadly. She is a former wildland firefighter and advocate for diversity in STEM. She likes to burn stuff.

Tea Lempiälä, Assistant Professor. Tea Lempiälä is an Asst Professor of Management at the Department of Management of Complex Systems. Her research interests include innovation practice and management, sustainable energy and paradoxical tensions of organizational life. Her research aims to create better understanding of innovation as a social and collaborative process, and through this discover better means for supporting it. Dr. Lempiälä is passionate about various aspects of equity, especially in the context of creating more inclusive innovation cultures. She has conducted qualitative research in technology companies and non-profit organizations in the United States, India and Europe.

Paul Maglio, Professor, Director of the Division of Management and Information.* Paul Maglio is a Professor of Management and Cognitive Science at the University of California, Merced. He holds a bachelor's degree in computer science and engineering from MIT and a Ph.D. in cognitive science from the University of California, San Diego. One of the founders of the field of service science, Dr Maglio is the Editor-in-Chief of INFORMS Service Science, and is lead editor of the Handbook of Service Science. He has published more than 125 papers in computer science, cognitive science, and service science.

Russ McBride, Assistant Professor.* Russ McBride did his Masters at Stanford and his PhD from UC Berkeley in philosophy and cognitive science. He then spent four years as a Research Professor at the University of Utah David Eccles Business School working on entrepreneurship and social theory, and working as the Director for The Foundry Entrepreneurship Incubator whose companies generated over $200M in revenue. His central research goal is to advance our understanding of social reality as it applies to firms, organizations, and entrepreneurship. Since, for social entities, what counts as a "real" social entity (like a firm, an academic department, or a government) depends upon what individuals believe is real (unlike physical objects which are real regardless of what we think), such research ultimately requires an exploration of aspects of human beliefs and cognition. He is the Director of the Social Reality & Cognition Research Group (SORAC: http://sorac.info). He has run the social ontology workshop at the Academy of Management Conference for the last 7 years. He has also spent many years doing research on artificial intelligence, running a software consulting company, and working on problems in cognitive science. Current works in progress include a book, "The Complexity of Human Behavior".

Tracey Osborne, Associate Professor, Presidential Chair in Management of Complex Systems.* Tracey Osborne is an Assoc Professor in the Department of Management of Complex Systems at the University of California, Merced. Her research focuses on the social and political economic dimensions of climate change mitigation in tropical forests and the role of Indigenous Peoples, the politics of climate finance (with particular emphasis on carbon markets), global environmental governance, and climate equity and justice. She has worked on these issues globally with extensive field experience in Mexico and the Amazon (Peru, Ecuador and Guyana). She also leads the Climate Alliance Mapping Project. This is a collaborative effort between academics, environmental non-governmental organizations, and Indigenous organizations working for a socially-just response to climate change through research, maps and digital stories. The mapping project is an initiative of the Public Political Ecology Lab, which Tracey founded and directs to support engaged scholarship by communicating environmental

research to broader publics. Her work has been published in high-impact geography and social science journals, and she has been invited to share her research internationally in academic and non-academic venues such as Conference of the Parties climate change meetings. She received her PhD from the Energy and Resources Group at the University of California, Berkeley.

Alex Petersen, Associate Professor.* Alex Petersen's research models the evolution of large multiscale socio-economic systems by applying concepts and methods from complex systems, statistical physics, management and innovation science. By way of example, recent empirical work analyzes two types of researcher mobility — cross-border and cross-disciplinary — by applying econometric and network science methods to measure career dynamics mediated by these relatively common career events. Other data-driven work addresses the growth of scientific production and its implications for collective attention, research evaluation, and mega-journal management.

LeRoy Westerling, Professor.* LeRoy Westerling is Professor of Management of Complex Systems and Chair of the Graduate Council of the Academic Senate at UC Merced. Prior to coming to Merced in 2006, he was a postgraduate research meteorologist and project scientist at Scripps Institution of Oceanography. He holds a BA from UCLA in International Economics and Chinese Area Studies, and a joint PhD in Economics and International Affairs from UCSD. Research interests include drivers of extreme fire events, wildfire climatology, and modeling and scenario analysis for wildfire and impacts affecting vegetation, carbon, water, air pollution and public health, habitat, insurance, and infrastructure.

Lisa Yeo, Assistant Professor.* Lisa Yeo's research focuses on people and process to help organizations design and build information systems that make it easy to protect privacy and prevent security breaches without requiring extensive investments in security layers after the fact. She explores topics of economics of information security and collaborative knowledge management through stochastic modelling, mathematical programming, and statistical analysis. Future research goals are designed to build on this foundation by introducing ideas and tools from behavioural operations into the decisions regarding information security, risk management, and collaborative knowledge production. Her work should inform policy and help design information systems that enable users to easily behave secure, privacy-protecting ways.

Courses

Teaching Plan and Faculty Workload

As mentioned, the MDSA program depends mainly on existing graduate courses. Core courses will be offered annually by the program's primary graduate groups – CIS, ECON, and MIST. Elective courses may be offered periodically by a larger set of graduate groups, including CIS, ECON, MIST, Environmental Systems (ES), Psychological Sciences (PSY), Political Science (PoliSci) and Public Health (PH). As a self-supporting program, MDSA must cover the costs of instruction from its own tuition revenues.

For courses provided by graduate groups other than the three core graduate groups of the program, we propose to share student fee revenues equally, just as UC Merced Extension shares concurrent enrollment revenues equally with academic programs. When appropriate, the MDSA program will create formal agreements with each of the other graduate groups (and their departments and schools) to allow placement of MDSA students into their courses in this way.

(see Appendix # for example MOU) MDSA students may be placed into existing courses as long as there is room, with priority given to state-supported graduate students. The current Graduate Group Chairs of Environmental Systems, Psychological Sciences, and Public Health have all expressed their support for this plan (see letters in Appendix G).

For courses provided by the primary MDSA graduate groups, students will be placed into existing sections of courses as long as there is room, with priority given to state-supported graduate students, under the same kind of revenue-sharing agreement under consideration with other groups. Once capacity limits are reached in existing sections of the core courses, a new section will be opened up and the instructor's time paid on an overload basis by the MDSA program (per APM 662). The anticipated number of total such overload sections based on the projected student enrollment is shown below (Table 7). We do not anticipate any additional sections to be needed for faculty members outside the program faculty, as the MDSA students will only be able to register for elective courses if there is sufficient space after ensuring priority is given to state-supported students. We expect to reach steady-state of 30 students by Year 4, requiring five additional sections for the core courses. If more than one extra section needs to be opened up, we will hire additional temporary faculty to teach in the program, with temporary instructional staff compensated on a per-class basis.

TABLE 7. STUDENT ENROLLMENT AND OVERLOAD SECTIONS, INCLUDING CAPSTONE

	Year 1	Year 2	Year 3	Year 4	Year 5
Student Enrollment	10	15	20	30	30
Overload Sections	1	1	6	6	6
Percent Overload	11.1%	11.1%	66.7%	66.7%	66.7%

Our plan is to rely almost entirely on regular graduate courses taught by ladder-rank faculty. Core courses will be offered each year. We expect electives will usually be offered every two or three years, but this depends on the schedule and demand of the course as determined by its home group. Of 46 core faculty across the groups, at least 6 will be involved in the delivery of core courses, and another 10 in the delivery of elective courses. As described no new courses (except for the capstone activity) will be needed, though additional sections of the core courses may be added when there is sufficient demand. The capstone activity will be advised by faculty who want to oversee student projects and internships, with the MDSA Program Chair serving as formal advisor if necessary.

We propose that the primary indicator for program sustainability is student enrollment: if the program fails to meet the anticipated enrollment targets based on a lagged two-year threshold, we should consider sunsetting the program. For example, we will consider sunsetting if there are no more than 10 students in Year 3 (corresponding to the anticipated number for Year 1), 15 students in Year 4, and 20 in Year 5. This plan will provide sufficient flexibility to accommodate uncertainties during the ramp-up phase.

Core Requirements and Courses

As described, there are six core requirements of the MDSA program. These may be fulfilled by specific courses from the program's primary graduate groups plus the completion of a capstone activity.

Methods of Data Science I

COGS 2XX: Data Science. Introduces the data analytics pipeline relevant to both academic and industrial work: obtaining raw unstructured data; cleaning, organizing, merging and identifying potential pitfalls in the data; exploring and visualizing the underlying statistics; managing data for publication and reproducibility. Introduces best-practices for handling and analyzing large multi-scale datasets using examples drawn from open-data repositories. Each instance of the course will include at least one detailed case study examining the integrated ethical and technical aspects of data science work. (Instructor: Dan Hicks)

Data Ethics

COGS XXX: Data Ethics. Introduces students to three kinds of work in data ethics: ethical analysis of existing and proposed data science products or processes, the ethical theories that justify and facilitate moral critiques, and the production of ethical evaluations of data science products and processes. Core concepts covered include privacy, big data collection & analysis, algorithmic bias, data ownership, and the social responsibilities of engineers. (Instructor: Hanna Gunn)

Statistics and Probability

COGS 210: Statistics. This course will introduce the statistical toolkit used by researchers in cognitive science to learn from data. The focus will be on (generalized) linear models (linear regression, logistic regression, multilevel models, etc.) from both frequentist and Bayesian perspectives. Students will learn to apply these techniques to real datasets using the R programming language, create publication-quality data visualizations, and build reproducible statistical workflows.. (Instructor: Rachel Ryskin)

Data Communication and Visualization

COGS 2XX. Global Good Lab. project-based graduate-level course with the goal of using data science to enact positive change in the world. At the beginning of the course, students will work in groups to identify issues impacting the world and will work backward to develop skills and knowledge to move the needle on the problems they've identified. At the end of the course, students will have produced 1-3 portfolio level projects (depending on the scope of the projects) and a communication plan for disseminating their work. Projects may include collaborations with industry and agency partners.(Instructor: Lace Padilla)

MIST 232: GIS Decision Analytics. Introduces geographic information systems (GIS) as the technology of processing spatial data, including input, storage and retrieval; manipulation and analysis; reporting and interpretation. Emphasizes GIS as a decision support tool for analysis, visualization, and problem solving in natural resource management, organizations/logistics, and non-profit sector. (Instructor: Jeff Jenkins)

Methods of Data Science II

ECON 2XX: Data Science for the Social Sciences. The course will mainly focus on two aspects of data science. First, it will address the theoretical and practical differences and complementarities between machine learning and causality estimation techniques. The benefits of classification trees or random forests will be compared to the ones of causal estimations like instrumental variable regressions. Second, text mining techniques like neural networks will be introduced in the context of real open-access datasets. The social and ethical implications of predictive analytics methods will be discussed. (Instructor: Christian Fons-Rosen)

Capstone Activity

MIST 2XX: Data Science and Analytics Integrative Project. Starting with an existing real-world problem identified with support of a project advisor (faculty, administration, industry), students will work in teams to propose, design, and develop a data-oriented solution. A viable portfolio product (e.g. interactive dashboard, entrepreneurial venture oriented around a data-service innovation, or data collection mobile app), report, and final presentation are required.

Themes

As described, students select one of three themes and must successfully complete three courses aligned with that theme. In the future, we expect to add more themes and more elective courses, potentially including courses from additional graduate groups.

Sustainability and Environment

COGS 222: Modeling Social Behavior. A broad approach to understanding and using formal models to study social behavior and organization. We will take an interdisciplinary approach, drawing from the social sciences and from evolutionary ecology in roughly equal measures. Methodology will focus on agent-based modeling, but also included some evolutionary game theory and related approaches. Topic covered include foraging, disease transmission, cooperation, segregation, social norms, and the cultural evolution of language. (Instructor: Paul Smaldino)

ES/MIST 232 Applied Climatology. Spatial and temporal patterns in climate and their association with land surface characteristics and processes. Methods for exploiting these for hypothesis testing, modeling, and forecasting. Applications include seasonal forecasting, ecological modeling, and analysis of processes such as flooding and wildfire. (Instructor: TBA)

ES 240: Water Resources Planning and Management. Basic concepts of and issues in water resources management, water resources planning, institutional and policy processes. Quantitative analytical methods in water resources planning and management; introduction to systems analysis, multi-objective planning, and risk assessment. Design project. Graduate requirements include preparation of a detailed case analysis. (Instructor: TBA)

ES 252: Remote Sensing of the Environment. Fundamental and advanced concepts of electromagnetic remote sensing, information extraction and applications in environmental monitoring. Advanced topics include principles of image extraction, image correction, image enhancement, classification methods, and new development of sensor

techniques. Reading materials and final research projects are required for graduate students. (Instructor: TBA)

ES 260: Sustainable Energy. Current systems for energy supply and use. Renewable energy resources, transport, storage, and transformation technologies. Technological opportunities for improving end-use energy efficiency. Recovery, sequestration, and disposal of greenhouse gases from fossil-fuel combustion. Graduate requirements include preparation of a detailed case analysis. (Instructor: TBA)

ES XXX: Life Cycle Assessment. Life cycle assessment (LCA) is a tool used across fields to determine the cradle-to-grave environmental impacts of products and systems. The course will cover how to mathematically define the life cycles of products and systems, perform an LCA, and interpret LCA results and evaluate them within the context of the scientific literature. Students in the course will individually conduct a complete life cycle assessment with a literature review, sensitivity analysis, and uncertainty analysis using available data and impact assessment methods with guidance from the instructor. (Instructor: TBA)

MIST 215: Political Ecology and Complexity. Provides background and tools for students to understand and engage with systems of natural resource governance through the complex relationships of political and ecological factors. Topics and case studies will be theoretical and applied in nature, and draw from literature in political ecology, complexity theory, and land use planning. Central to this course will be the competing knowledge claims, power structures, and values that policy-makers, scientists, and the public deal in. (Instructor: Jeff Jenkins)

MIST 254: Systems Thinking for a Just and Sustainable Future. Introduces systems thinking, key concepts and methods, particularly related to climate change and sustainability. Systems thinking is based on relationships, patterns and context. This way of thinking is crucial for understanding the complexity of major social and ecological problems, and determining leverage points for effective solutions. Students will develop the analytical tools for systems thinking and be better equipped to make decisions that support a socially just and environmentally sustainable future. (Instructor: Tracey Osborne)

PH 204: Environmental Health. Why the environment is essential to human health and how we analyze and act on environmental agents, factors, and conditions to improve health of people, at local, regional, and global scales. Focuses on issues pertinent to the San Joaquin Valley. (Instructor: Ricardo Cisneros or Asa Bradman)

Human Behavior

COGS 204: Complex Adaptive Systems. Introduction to the study of complex phenomena using dynamical computer simulations, which exhibit emergent properties, sensitivity to initial conditions, fractal structure, phase transitions in random graphs, and shifts from stability to meta-stability to chaos. Matlab projects include probability games, neural networks, the Lorenz attractor, the logistic map, the Mandelbrot set. (Instructor: Michael Spivey)

COGS 222: Modeling Social Behavior. A broad approach to understanding and using formal models to study social behavior and organization. We will take an interdisciplinary approach, drawing from the social sciences and from evolutionary ecology in roughly equal measures. Methodology will focus on agent-based modeling, but also included some evolutionary game theory and related approaches. Topic covered include foraging, disease transmission,

cooperation, segregation, social norms, and the cultural evolution of language. (Instructor: Paul Smaldino)

ECON 205: Applied Econometrics. This course teaches students the essential toolkit of causal analysis. Whenever we're interested in understanding the effect of government action or business policy or individual choice, we must use a research design to distinguish correlation from causation: What is the effect of minimum-wage laws on employment? Pension incentives on productivity? An additional year of education on earnings? These are some of the questions we learn to tackle in class. This course presents the most advanced techniques for addressing causal problems by identifying and exploiting natural settings that approximate randomized trials. The course also prepares graduate students for professional life, instructing programming, visual presentation, and writing. (Instructor: Andrew Johnston)

MIST 254: Systems Thinking for a Just and Sustainable Future. Introduces systems thinking, key concepts and methods, particularly related to climate change and sustainability. Systems thinking is based on relationships, patterns and context. This way of thinking is crucial for understanding the complexity of major social and ecological problems, and determining leverage points for effective solutions. Students will develop the analytical tools for systems thinking and be better equipped to make decisions that support a socially just and environmentally sustainable future. (Instructor: Tracey Osborne)

MIST 261: Qualitative Methods for Management and Innovation Studies. Provides an understanding of how to conduct qualitative research particularly in the context of management, innovation and sustainability studies. The course offers an understanding of the major philosophical and epistemological underpinnings of qualitative research strategies and introduces the students to a range of qualitative research methodologies. Students learn how to define qualitative research questions, design a qualitative study, and write up qualitative research through examples from existing research and exploration of their own research ideas. (Instructor: Tea Lempiala)

MIST 271: Network Science. Introduces the theory of network science with a focus on real-world applications. Lectures and discussion engage students around examples drawn from scientific literature that emphasize the importance of identifying, understanding, and measuring complex network structure. Mathematical theory will be paired with real-world examples of complex networks encountered in biology, environment, society, and business. Computational lab will demo software for analyzing and visualizing networks that are constructed from real-world data. A team project culminating in a final presentation will provide the opportunity for students to apply course material to their ongoing research. (Instructor: Alex Petersen)

MIST 2XX: Service Innovation. Focuses on information-oriented service innovation and generation of new successful IT/data-oriented service ventures in the digital economy. Helps students gain the skills necessary to be successful in three main aspects of web-mediated service production and delivery systems: the backend, the front-end, and service design. (Instructor: Paul Maglio)

POLI 213: Experimental Methods in Political Science. This course is intended to provide students with an understanding of experimental methods in Political Science. The first portion of the semester will emphasize concepts and tools from the experimentalist's toolbox with a strong focus on causal inference, external and internal validity, and choosing subjects and subject's motivations. Later weeks will focus on issues and challenges to specific types of Political

Science experiments including survey experiments, laboratory experiments, physiological experiments, and field experiments. (Instructor: Elaine Denny)

POLI 219: Behavioral Game Theory - Special Topics in Political Science Methodology. Behavioral game theory uses a variety of methods to understand the connection between game theory and human behavior. Game theory is the predominate method for mathematically modeling and analyzing strategic interactions in political science and economics. However, many conventional modeling assumptions are based on introspection and guesses, rather than careful observation of how people actually play in games" (Camerer 2003, pg. 3). Behavioral game theory tries to figure out when these simplifying assumptions provide insight into human behavior. It also tries to develop models that can explain behavioral deviations from the traditional assumptions and tries to test the additional hypotheses that these models generate. Throughout this course, students will learn how experiments and other empirical methods have been used to answer important questions related to game theory and game theoretic models. Doing so will hopefully give students a better understanding of how you might use similar experimental tools in your own research. (Instructor: Brad LeVeck)

POLI 251: Political Cognition. The seminar provides students with an overview of political psychology. How do we think and feel about politics? What preferences do we have and what motivates us to become engaged? How does the political context shape our political psychology? And how can we answer these questions? We will draw on research from political science, including the study of American politics, comparative politics, and international relations, as well as cognate disciplines like economics, psychology, sociology to answer these and other questions. (Instructor: Christopher Ojeda)

Policy and Decision-Making

COGS 222: Modeling Social Behavior. A broad approach to understanding and using formal models to study social behavior and organization. We will take an interdisciplinary approach, drawing from the social sciences and from evolutionary ecology in roughly equal measures. Methodology will focus on agent-based modeling, but also included some evolutionary game theory and related approaches. Topic covered include foraging, disease transmission, cooperation, segregation, social norms, and the cultural evolution of language. (Instructor: Paul Smaldino)

ECON 205: Applied Econometrics. This course teaches students the essential toolkit of causal analysis. Whenever we're interested in understanding the effect of government action or business policy or individual choice, we must use a research design to distinguish correlation from causation: What is the effect of minimum-wage laws on employment? Pension incentives on productivity? An additional year of education on earnings? These are some of the questions we learn to tackle in class. This course presents the most advanced techniques for addressing causal problems by identifying and exploiting natural settings that approximate randomized trials. The course also prepares graduate students for professional life, instructing programming, visual presentation, and writing. (Instructor: Andrew Johnston)

MIST 254: Systems Thinking for a Just and Sustainable Future. Introduces systems thinking, key concepts and methods, particularly related to climate change and sustainability. Systems thinking

is based on relationships, patterns and context. This way of thinking is crucial for understanding the complexity of major social and ecological problems, and determining leverage points for effective solutions. Students will develop the analytical tools for systems thinking and be better equipped to make decisions that support a socially just and environmentally sustainable future. (Instructor: Tracey Osborne)

PH 202: Epidemiological Methods. Introduces, compares, and applies conceptual frameworks, measures, study designs, and analysis approaches used in the field of epidemiology including causality, measures of disease, measures of association, study design (trials, cohort, case-control, cross-sectional and ecological), biases, screening, statistical inference, and analyzing epidemiologic data. (Instructor: Sidra J. Goldman-Mellor, Sandie Ha, or Irene Yen)

PH 204: Environmental Health. Why the environment is essential to human health and how we analyze and act on environmental agents, factors, and conditions to improve health of people, at local, regional, and global scales. Focuses on issues pertinent to the San Joaquin Valley. (Instructor: Ricardo Cisneros or Asa Bradman)

PH 216: Health Policy. Examines key health policy research topics and methods used in the field. The course is designed to provide students with an overview of health policy issues and research in the field while simultaneously developing rigorous critical analysis and research skills. Students will learn about a range of local, state, and federal policies in the U.S. to improve population health outcomes, including health care policies (e.g., Medicare, Medicaid, etc.) and public health policies to promote health. They will also learn about the health policymaking process. Research articles in leading peer-reviewed journals in the fields of health policy, public health, health services research, economics, and political science will be emphasized. (Instructor: Denise Payan)

PSY 224: Health Disparities. Disease prevalence, severity, and treatment varies across sociodemographic groups. Understanding why health disparities occur is key to determining how inequalities might be alleviated. The focus of this course is on research that a) describes health disparities, b) investigates factors that explain differences, and c) proposes interventions to treat at-risk populations. (Instructor: Anna Epperson)

PSY 225: Health Risk Decision Making. A focus on the decision making process underlying health risk behaviors. Consideration of the role perceptions of risks/benefits, attitudes, emotions, social relationships, and the media play on health decisions, with an emphasis on decision making theories (e.g., rational choice theory, prospect theory, health beliefs model, and the theory of planned behavior). (Instructor: Anna Song)

POLI 219: Model Based Inference - Special Topics in Political Science Methodology. This is an advanced graduate methodology class which will focus on the use of empirical models for statistical description and prediction with social science applications. The class takes a "learning" approach to modeling, covering GLMs and machine learning. this class follows from the probability, regression, and causal inference classes in the political science methodology sequence. (Instructor: Tesalia Rizzo)

POLI 224: Subnational Politics. This seminar provides an overview of major debates in research on subnational politics. We examine origin and evolution of subnational political environments. We trace the historical development of local government institutions, analyze coalitions, investigate distributions of power, investigate the process and consequences of suburbanization,

and examine local development and education policy. The primary goals of the course are to familiarize students with the principal questions asked by scholars in this subfield, the methodological approaches employed, and the avenues available for future research. Methodological diversity will be emphasized. (Instructor: Jessica Trounstine)

Data Bootcamp

In collaboration with the UC Merced Library, the MDSA program will host a multi-day *Data Bootcamp* before the start of Fall semester for students who require grounding in concepts and tools of data science. The bootcamp will incorporate workshops in data literacy and computer literacy, introduction to programming in R and in Python, working with spatial data, and cleaning data.

Day 1 – Data Literacy/Computer Literacy

Introduction to Locating Data Sets & Data Repositories

Students will

- explore major data sources, directories and repositories.
- become familiar with the different levels of access to data sets (open vs. licensed).
- locate discipline specific repositories and in-depth training resources.

Locating research data is essential to research reproducibility, facilitating collaboration, and raising new insights that come from secondary analysis. However, as a researcher, how do you locate data sets that are relevant to your research and contain reliable data? This workshop will provide an introduction to the data repository landscape and strategies for navigating data sources to discover relevant and usable data.

Open Data & Creative Commons

Students will

- articulate the concept of open data and its value to research reproducibility.
- understand how Creative Commons licenses are used to share open data.
- learn how to apply and select Creative Commons licenses for the data that you create and wish to share.
- become familiar with large, openly available data sets and their licensing.

There exists a vast array of data sets and other resources that is freely available to the public for use in research. While some of these resources are available for use without permission, many of these data sets and resources are made available under Creative Commons licensing. Creative Commons is a nonprofit that offers simple, standardized open licensing that attaches copyright permissions to resources and ensures that proper author(s) attribution is given, so that scholars may use those resources to advance their own research. This workshop will examine the concept of open data and resources, explain Creative Commons licensing, and review large, openly available data sets and their licensing.

Data Use & Ethics: Copyright and Licensing (50 min.)

Students will

- recognize key issues in e-resource licensing.
- understand the roles of copyright and licenses.
- distinguish between what is ethical and what is legal.
- determine what user behavior is acceptable/permitted under license terms.

Many information sources, at universities, organizations and corporations, are made available to users under specific licensing conditions. This session will introduce key concepts related to copyright and licenses and highlight desirable licensing terms for e-resources subscribed to at the UC Merced Library. Participants will view sample licenses and determine acceptable/unacceptable user behavior under the terms provided.

Introduction to the Command Line

This workshop teaches the basics of the Unix command line, including:

- Working with files and directories
- File and process redirection
- Composing custom workflows with command-line tools
- Project organization

Day 2 – Introduction to Programming in R and Python

A brief introduction to Python

This workshop teaches the fundamentals of research computing with Python, including:

- Built-in data types and functions
- Using program control flow to write custom scripts
- Working with files
- Debugging and finding help

A brief introduction to R (3 hours)

This workshop teaches the fundamentals of research computing with R, including:

- Built-in data types and functions
- Using program control flow to write custom scripts
- Working with files
- Debugging and finding help

DAY 3 – WORKING WITH SPATIAL DATA AND CLEANING DATA

INTRODUCTION TO ARCGIS (2 HOURS)

Introduces the GIS interface, basic GIS concepts, data management best practices and examples on how GIS is used in the data science field.

CLEANING MESSY DATA

Students will

- recognize the value of Open Refine for consistent and efficient data practices.
- use OpenRefine functionality to clean and organize data (facet, cluster, split, transform).
- export cleaned data into other formats.
- save cleaning scripts for re-use.

OpenRefine is a free open source tool designed specifically for cleaning and organizing your data with easy to use features for fixing errors and transforming your data. Learn more about how this powerful tool can help you reduce the time spent cleaning and prepping your data as you prepare your data for analysis and sharing.

Resource Requirements

The draft budget for the self-supporting graduate degree program, Master of Data Science and Analytics, sets the program fee for students at $45,000, returns 20% to student aid, and a 30% assessment for the campus. As a self-supporting graduate degree program, our Master of Data Science and Analytics will pay for itself and generate income for use by the school, as well as committing substantial resources in return-to-aid to support a diverse student body and enhanced opportunities for California students. The draft budget plan shows the program operating at a small loss in the pre-opening year (see Appendix A). Any such loss will be covered by endowed funds available to the proposed Gallo School and will be repaid as the MDSA becomes financially viable over the first two years of operation.

Faculty FTE

The proposed MDSA program is planned to operate with existing faculty resources, requiring no new faculty FTE. As described, students will be placed into existing courses (by arrangement with related graduate groups and programs) or additional sections of existing courses will be offered on an overload-teaching basis. If multiple sections need to be added, additional faculty (either ladder-rank, temporary instructors, or graduate student instructors) will be paid to teach them. All costs of instruction will be paid from the self-supporting program revenues.

Staff FTE

The proposed MDSA program plans to add dedicated staff to help manage the program over time – starting with a half-time position in Year 1 and growing to one full time position by Year 5. Salary and benefits will be paid by MDSA program revenues. No specific staff resources from the campus are required.

TA Support

As described, the proposed MDSA program will support TAs for dedicated sections of core courses from its own revenues. No additional TA support from the campus is required.

Space and Other Capital Facilities

Given a shortage of computer teaching labs on campus – with existing labs used more than 95% of the time today – we expect increased demand from data science students may require increased teaching lab capacity over time, which may be defrayed by support from the SSGPDP in Data Science and Analytics. Alternatively, it may be appropriate to deliver lab portions of the data science curriculum virtually, limiting the need for additional on-campus resources.

Equipment

We do not anticipate any new equipment cost during the first five years, but if any need arises in the future, it will be fully paid by the program.

Computing Costs

We anticipate that all students in the program will have their own laptop computers and most of the coursework will be performed using these. However, the program will purchase new computers ($3,000/unit) for the students who demonstrate financial need as described in the *Graduate Student Support* section below.

Library Acquisitions and Cost

Though we do not anticipate any impact on collections and though 30% of program revenues will be returned to central campus for overhead (including presumably some amount for the library), overall support from the library (e.g., SpARC's support for GIS applications during the year, etc) is critical. Our support for the library comes in two parts:

> **Operations Support**: MDSA will provide $5,000 annually for operational support.
>
> **Workshop Support**: MDSA will provide support per student enrolled in the proposed data bootcamp facilitated by the library (estimated to be about $333/student). The MDSA and bootcamp program will start with 10 students, potentially growing to 40 over 5 years. In later years (>25 students), we may require multiple sections of the bootcamp workshops and we will adjust the fee structure as needed.

Other Operating Expenses

The full budget is provided in Appendix A and summarized in Table 8.

TABLE 8. MAJOR OPERATING COSTS

Expense	Year 1	Year 2	Year 3
Instructional Budget	$97,500	$142,825	$176,291
TAs (incl fees)	-	$64,835	$66,780
Director Stipend	$30,900	$31,827	$32,782
Staff (Salary/Benefits)	$48,653	$50,112	$51,615
Supplies/Expenses	$85,000	$88,100	$96,310
Student Aid	$90,000	$140,400	$194,688

Summary of Projected Budgetary Needs

None.

Graduate Student Support

Data science masters programs typically prepare graduates for careers in the technology sector, and so it is not standard to provide students direct financial support. However, by design our proposed MDSA program targets students with a more diverse disciplinary pedigree than traditional data science programs, and consequently we anticipate students will use this one-year program as a launching pad not only for careers in technology, but also for other informatics-oriented professions that do not necessarily feature high starting salaries. In addition, we anticipate drawing largely from the same population as UC Merced's undergraduates, contributing to the development of a diverse informatics workforce, and requiring substantial financial support. Thus, the MDSA program will return at least 20% of revenues to student aid, which is a far higher proportion of return-to-aid than any other SSGPDP program we have found in the UC system.

To deliver student aid, we will establish a need-based scholarship program to help students offset program costs, living expenses, and the costs of supplies and equipment. For instance, to ensure all enrolled students have a personal computer that is appropriate for our computationally intensive program, we will provide students with financial need with up-to-date computing hardware, such as an Apple Macbook Pro laptop with multi-core processor for parallel computing, 16GB RAM for loading large datasets, and a second portable monitor, totaling roughly $3,000 per qualifying student. In addition, the capstone project may be another source of possible student support, for instance in the case that a student matches with a faculty or other external sponsor that has funding to compensate the capstone team for the timely delivery of a viable product at the end of the project.

Equity, Diversity, and Inclusion

Like most of the technology industry, data science is not particularly diverse. A 2020 industry survey of 1,001 data scientists found that 71% were men.[61] A 2017 analysis of part-time students in general data science courses found that 46% were White and 28% were Asian, while only 8% Latinx and 4% African American.[62] In addition, critical race scholars have noted that a broad "Asian" racial category elides important class and cultural differences between, for example, born citizens, white-collar professional immigrants, and refugees. Scholars of race and technology have argued that the underrepresentation of women and people of color in the tech industry has contributed to problems such as embedding racist and sexist stereotypes into search results.[63]

The proposed MDSA program at UC Merced is poised to help address this underrepresentation problem in data science two ways. First, the program is designed to be open to students with broad undergraduate backgrounds, not just those with a degree in computer science, engineering,

[61] See https://365datascience.com/career-advice/career-guides/become-data-scientist-2020/

[62] See https://generalassemb.ly/blog/data-science-gender-race-disparity/

[63] Noble, Safiya Umoja. 2018. Algorithms of Oppression: How Search Engines Reinforce Racism. NYU Press.

or mathematics. Members of underrepresented groups who majored in fields such as biology (15% of 2019-2020 US Bachelor's degrees awarded to Latinx students; 9% to Black students) or social science (19% Latinx; 10% Black) will be able to use the program as a gateway to the tech industry.[64] Second, by targeting the same student population for the proposed program as UC Merced's undergraduate programs, we expect to draw from a very diverse pool: Over the last five years, UC Merced has awarded bachelors' degrees to 50% Hispanic, 5% African American, 22% Asian, and 12% White students.[65]

However, the relationship between the technology industry and social injustices is not just a matter of underrepresentation. Structural factors and high-level business decisions are also substantial. For example, Ring, Inc., is an Amazon subsidiary that manufactures and sells home surveillance technology, particularly web-connected doorbell cameras. Ring has actively cultivated relationships with law enforcement and has provided advice on how law enforcement can request access to recordings from Ring devices.[66] Take another example: A recent study found the LGBTQ professionals in STEM fields, such as data science and data analytics, are more likely to experience career limitations and harassment, reported more frequent health difficulties, and were more likely to intend to leave the field than non-LGBTQ peers – and that these depend mainly on workplace factors.[67] Simply diversifying the data science workforce will do little to address both such behaviors. In this context, our Ethics PLO is crucial, as it includes both training in ethical theory and techniques of applied ethical analysis as well as character formation. Students will be taught not only how to ethically evaluate projects, but also to refuse to participate in projects that are severely and unavoidably unethical.

The creation of the proposed MDSA degree presents an opportunity to address these equity, diversity, and inclusion issues in data science. A commitment to diversity in faculty, students, and research populations is an essential element of the current practices within each group and department collaborating on the MDSA program. More specifically, MDSA will advance UC's goals for diversity, equity, and inclusion in a number of ways:

- Create new pathways to ensure both **recruitment and retention of underrepresented students**, including LGBTQ, non-binary, and disabled students.
    - In describing the program, we will emphasize how inclusion will be prioritized, and at orientation, we will discuss the importance of dignity and upholding the value of difference.
    - Faculty will be required to undergo specific EDI training, and the CDO has agreed to provide inclusive pedagogy seminars for our program.

---

[64] Data available at https://nces.ed.gov/programs/digest/d21/tables/dt21_322.30.asp

[65] Data available at https://visualizedata.ucop.edu/t/UCMerced/views/DegreesAwarded-Internal/DegreesAwardedStory/a5c4b06d-ba7b-41c9-a719-a7a0f5ebe47d/0d5d5ce3-ccbf-4ba9-aba0-2bc3b75397e9?:display_count=n&:showVizHome=n&:origin=viz_share_link

[66] See https://www.washingtonpost.com/technology/2019/08/28/doorbell-camera-firm-ring-has-partnered-with-police-forces-extending-surveillance-reach/, https://www.vice.com/en_us/article/43kga3/amazon-is-coaching-cops-on-how-to-obtain-surveillance-footage-without-a-warrant, and https://www.wired.com/story/ces-2020-amazon-defends-ring-police-partnerships/.

[67] Cech, E. A. & Waidzunas, T.J. (2021). Systemic inequalities for LGBTQ professionals in STEM. Science Advances, 7(3), 10.1126/sciadv.abe0933.

- We will ensure that every course syllabus includes a statement of how inclusion and upholding dignity are critical and provides resources describing how underrepresented students can connect to campus offices for allyship, support, or celebration of their identities.

- We understand that no one person can be characterized by a single dimension, and we recognize that categories, such as first-generation and working class or racial and ethnic membership and social class, cooccur. Our goal is to create a new model of human interaction within our program to capitalize on the bringing together of different lived experiences and perspectives.

- We will aim to ensure students see themselves reflected in relevant industries and positions (separate from the faculty and campus). For courses with guest speakers, we will ensure at least one guest speaker is a member of underrepresented group. For recruiting activities, we will target at least some companies that are minority-led or women-led.

- Program resources will be allocated for recruiting trips to undergraduate programs enrolling high numbers of underrepresented students, including California State Universities.

- Using at least 20% of revenues for student aid, we will establish fellowships for underrepresented students with demonstrated financial need; as program revenues increase, the number of fellowships will increase.

- Retention of underrepresented students (within the one-year program) will be enhanced through targeted mentoring that incorporates faculty and others from underrepresented groups, and on program staff to provide targeted job/internship support that is tailored to the needs of individual students.

- Create a student advisory group to serve as a conduit to report issues that students might be having in the program.

- Integrate **social justice and ethics** into the curriculum.

  - The curriculum includes a full course on Data Ethics, and ethical topics will be integrated as modules into other courses. (For an example, see the sample syllabus for Methods of Data Science I, Appendix D.)

  - The assessment rubric (Appendix B) covers both theoretical knowledge of the ethical, legal, and social implications (ELSI) of technology as well as the application of this theory to data science work.

- Theme courses may also touch on specific ethical and social justice issues in other contexts, e.g., environmental justice and courses in the *Sustainability and Environment* theme.

- Ensure **leadership opportunities** for faculty from minoritized groups.

  - To establish a leadership pipeline from faculty across the constituent groups, we will nominate and support their participation in professional development, such as the Inclusive Excellence Institute that will be sponsored by the Vice Chancellor/Chief Diversity Officer, and professional development opportunities for faculty to participate in national programs that provide leadership development.

- Develop **community** for faculty and students from minoritized group.

  - We will create opportunities for our alumni to be continuously engaged in the data science program, cultivating a diverse culture of inclusive learning and student-to-alumni connections.

  - We will seek to develop research funding opportunities for underrepresented faculty members and students through targeted development efforts in conjunction with the campus's Office of Philanthropy and Strategic Partnerships.

  - We will host periodic workshops or conferences focused on diversifying the data science and analytics profession, inviting professional, faculty, and students to discuss challenges and innovative approaches to increasing diversity.

To evaluate our progress, we will track the number of faculty and students from historically minoritized groups and track the number of these faculty in leadership roles. We will also track the job and internship placement of minoritized students. To track our performance and outcomes, we will consider all aspects of diversity, including race, ethnicity, gender identity, among others, and we will administer annual surveys for students and faculty to provide anonymous feedback on culture and climate in the program, aiming to capture whether faculty and students feel respected and appreciated.

We will continually assess climate, workload, and satisfaction among both students and faculty in the program. We will be very intentional in capturing demographic information without identifying individuals, and we will be very sensitive in how we report the data given small sample size, especially as we start. Our goal is to ensure that we are promoting perspective taking and a growth mindset, and that dignity and respect characterize the classroom in our program. Program faculty will come together on a regular basis to review the data with the commitment to continuous improvement. In addition, the Program Director will meet with the Associate Dean for Equity, Justice and Inclusive Excellence, and the Executive Committee of the proposed Gallo School (or of the School of Engineering, as appropriate) at least annually to review the program's strategic diversity plan and make necessary revisions to align itself with the campus priorities and system-wide initiatives.

Governance

The Master of Data Science and Analytics Program will be overseen by a Program Director and an Executive Committee whose membership includes one faculty member from the three affiliated graduate groups of the Program—CIS, ECON and MCS. The program will be governed by this Executive Committee rather than a graduate group because the proposed MDSA program is a self-contained one-year SSGPDP based upon pre-existing curriculum. Duties of the Program Director and Executive Committee are laid out in the proposed program bylaws (see Appendix I). Core members of the affiliated Graduate Groups are automatically Program faculty members and will vote on major decisions related to the Program when prompted by the Executive Committee. The Program Director is appointed by the appropriate Dean. The Executive Committee will solicit the names of nominees from Program faculty, submit the nominees' names for comments by Program faculty, and provide at least two names of nominees to the Dean. The Program Director serves as the Head Graduate Advisor, and chairs the Executive Committee, which makes decisions about hiring instructors and staff, and appoints or acts as the Program's Membership Committee, Admissions Committee, and Education Policy Committee.

Duties of the MDSA Program Director include marketing, recruitment, professional engagement to link students with capstone projects involving faculty/industry partners, managing curriculum and course schedules, serving as formal adviser to all MDSA students, program assessment and review, and managing all finances of the self-supporting operation (including collecting revenues, paying for instruction, and maintaining an up-to-date budget). To ensure the program director's incentives are aligned with program, school, and campus needs, we will create a contract for the program director that ties salary increases (projected at 3% annually) and contract renewal to program performance, particularly program finances and program enrollment (details of this agreement will be developed by the appropriate dean in consultation with the Executive Committee).

Program review will be conducted by the Program Review Oversight Committee (PROC), as is the case with all graduate programs, following the schedule established by PROC and the guidelines set in CCGA's Suggested Guidelines for Periodic Academic Review of SSGPDPs. The Program Director shall work with the Executive Committee to prepare review materials. In addition, our Industrial Advisory Board will provide input on the program review.

Changes in Senate Regulations

None required.

Appendices

Appendix A – Budget

Master of Data Sciernce and Analytics							
		2023-24	2024-25	2025-26	2026-27	2007-28	2028-29
FTEs:		Pre-opening					
MS Student Enrollment			10	15	20	30	30
Teaching Faculty (# Courses)			8	8	8	8	8
Lecturer FTE			0	0	0	0	0
25% TA/Year			0	3	3	4	5
Staff FTE		0.25	0.5	0.5	0.5	1	1
Operating Revenues:							
Annual Total Student Fees			50,070	52,073	54,156	56,322	58,575
Program Fee			45,000	46,800	48,672	50,619	52,644
Campus Fee			1,550	1,612	1,676	1,743	1,813
Health Insurance			3,521	3,661	3,808	3,960	4,118
Total Operating Revenues			**450,000**	**702,000**	**973,440**	**1,518,566**	**1,579,309**
Operating Expenses:							
Instructional S&B							
Faculty Teaching	varies		97,500	142,825	176,291	215,974	223,466
Addtl faculty	-	-	-	-	-	-	
TA Salaries	11,275	-	-	34,839	35,884	49,281	63,449
TA Benefits	*5%*	-	-	1,742	1,794	2,464	3,172
TA Tuition	8,619	-	-	28,254	29,102	39,966	51,457
Overload teaching				-	-	-	-
Overload benefits				-	-	-	-
Lecturers	*75,000*	-	-	-	-	-	-
Lecturer Benefits	*5%*	-	-	-	-	-	-
Total Instructional S&B		**-**	**97,500**	**207,660**	**243,070**	**307,685**	**341,544**
Non-Instructional S&B							
Director Stipend	*30,000*	15,000	30,900	31,827	32,782	33,765	34,778
Director Benefits	*3%*	450	927	955	983	1,013	1,043
Staff Salaries	*65,000*	15,000	32,500	33,475	34,479	71,027	73,158
Staff Benefits	*49.7%*	7,455	16,153	16,637	17,136	35,301	36,360
Total Non-Inst S&B		**37,905**	**80,480**	**82,894**	**85,381**	**141,106**	**145,339**
Non-Personnel Expenses							
Infrastructure		-	10,000	10,300	10,609	10,927	11,255
Supplies/Administrative		5,000	15,000	15,450	15,914	16,391	16,883
Supplies/Instructional		7,500	15,000	16,500	18,150	19,965	21,962
		-		-	-	-	-
Workshop/Orientation		-	10,000	10,500	11,025	11,576	12,155
Marketing/Advertising	*10.0%*	30,000	36,000	56,160	77,875	121,485	126,345
Faculty Research Support			-	-	-	10,000	10,000
Graduate Group Support	*1.0%*		-	-	9,734	15,186	15,793
Student Aid (IDC exempt)	*20.0%*	-	90,000	140,400	194,688	303,713	315,862
Total Non-Personnel Costs		**42,500**	**176,000**	**249,310**	**337,995**	**509,244**	**530,254**
Indirect Costs							
Campus Assessment Fees	*30.0%*	-	79,194	119,839	141,527	196,297	210,383
Total Indirect Costs		**-**	**79,194**	**119,839**	**141,527**	**196,297**	**210,383**
Total Operating Expenses		**80,405**	**433,173**	**659,703**	**807,974**	**1,154,331**	**1,227,519**
Annual Costs Per Student							
Program Direct Costs		-	26,398	26,631	23,588	21,811	23,376
Return to Aid			9,000	9,360	9,734	10,124	10,529
Indirect Cost			7,919	7,989	7,076	6,543	7,013
Total Costs Per Student			**43,317**	**43,980**	**40,399**	**38,478**	**40,917**
Profit (Loss) Per Student		**-**	**1,683**	**2,820**	**8,273**	**12,141**	**11,726**
Program Surplus (Deficit)		(80,405)	16,827	42,297	165,466	364,235	351,790
Program Cumulative Balances		(80,405)	(63,578)	(21,281)	144,185	508,420	860,210

Cost Analysis for Self-Supporting Program Fee Proposals		2022-23	Department-based SSGPDP	
Location:	Merced On-Campus	Program:	Data Science & Analytics	
line #		2024-25 Est.	2025-26 Proj.	2026-27 Proj.
	FTE ENROLLMENT			
1	Year-average Program Enrollment (FTE)	10	15	20
2	Year-average Program Headcount	10	15	20
	REVENUE			
3	Annual Fee Per Student (Fee detail is optional)	$50,070	$52,073	$54,156
	Program Fee	$45,000	$46,800	$48,672
	Student Services Fee			
	Campus-Based Fees (SU+FL+SP)	$1,550	$1,612	$1,676
	Health Insurance	$3,521	$3,661	$3,808
	Other (Example: Transit Fees)			
4	Total Fee Revenue Generated	$450,000	$702,000	$973,440
5	Total Other Funds (describe donation for student award)	$0	$0	$0
6	**TOTAL PROGRAM REVENUE**	***$450,000***	***$702,000***	***$973,440***
	COSTS			
	A. Program Direct Costs, Subject to IDC			
7	Total Faculty Salaries	$97,500	$140,250	$171,870
8	Total Staff Salaries	$63,400	$65,302	$67,261
9	Faculty and Staff Benefits	$17,080	$20,167	$22,540
10	General Assistance	$0	$64,835	$66,780
11	S&E	$86,000	$108,910	$133,573
12	Equipment	$0	$0	$0
13	Travel	$0	$0	$0
14	Campus-based fee-funded activities (if any)	$0	$0	$0
15	Other (describe ___graduate group support _____)	$0	$0	$9,734
16	***TOTAL DIRECT COSTS, SUBJECT TO IDC***	***$263,980***	***$399,464***	***$471,758***
	B. Program Direct Costs, Exempt from IDC			
17	Financial Aid	$90,000	$140,400	$194,688
18	Other S&E (describe _____)	$0	$0	$0
19	Other Equipment (describe _____	$0	$0	$0
20	Other (describe *Ex: awards to students*	$0	$0	$0
21	***TOTAL DIRECT COSTS, EXEMPT FROM IDC***	***$90,000***	***$140,400***	***$194,688***
	C. Total Direct Costs			
22	***TOTAL PROGRAM DIRECT COSTS (line 16 + line 21)***	***$353,980***	***$539,864***	***$666,446***
	ANNUAL COST PER FTE STUDENT			
23	Program Direct Costs *(line 22 / line 1)*	$35,398	$35,991	$33,322
24	Program IDC Rate	30.00%	30.00%	30.00%
25	Program Indirect Costs *([line 24 x line 16] / line 1)*	$7,919	$7,989	$7,076
26	**TOTAL COST PER FTE STUDENT**	***$43,317***	***$43,980***	***$40,399***
27	**TOTAL PROGRAM COST** *(line 1 x line 26)*	***$433,173***	***$659,707***	***$807,980***
28	**SURPLUS (DEFICIT)** *(line 6 minus line 27)*	***$16,827***	***$42,293***	***$165,460***
29	**SURPLUS (DEFICIT) PER HEADCOUNT STUDENT**	**$1,683**	**$2,820**	**$8,273**

Attachment 3: Cost Analysis

Enrollment - Part 2

Campus: Merced
Program: Data Science & Analytics

YEAR-AVERAGE ENROLLMENTS

		2024-25 Est.	2025-26 Proj.	2026-27 Proj.
	Self-supporting FTE Enrollments*			
1	Master of Data Science & Analytics	10	15	20
2	**Total Program Enrollment**	10	15	20
	Year-Average Headcount Students - This Program			
3	class/cohort 1	10	15	20
4	class/cohort 2	0	0	0
5	class/cohort 3	0	0	0
6	class/cohort 4	0	0	0
7	class/cohort 5	0	0	0
8	class/cohort 6	0	0	0
9	class/cohort 7	0	0	0
10	class/cohort 8	0	0	0
11	Total	10	15	20
12	Conversion Ratio	**1.0000**	1.0000	1.0000
	Fees by Class - This Program			
13	class/cohort 1	$45,000	$46,800	$48,672
14	class/cohort 2	$0	$0	$0
15	class/cohort 3	$0	$0	$0
16	class/cohort 4	$0	$0	$0
17	class/cohort 5	$0	$0	$0
18	class/cohort 6	$0	$0	$0
19	class/cohort 7	$0	$0	$0
20	class/cohort 8	$0	$0	$0
	Estimated Fee Revenue			
21	class/cohort 1	$450,000	$702,000	$973,440
22	class/cohort 2	$0	$0	$0
23	class/cohort 3	$0	$0	$0
24	class/cohort 4	$0	$0	$0
25	class/cohort 5	$0	$0	$0
26	class/cohort 6	$0	$0	$0
27	class/cohort 7	$0	$0	$0
28	class/cohort 8	$0	$0	$0
29	**Total Fee Revenue**	**$450,000**	**$702,000**	**$973,440**

Appendix B – Student Assessment Rubric

PLO	Expert	Advanced	Intermediate	Introduced
Methods	(a) Works fluently with both general methods and tools and some set of specialized methods and tools in at least one core programming language (R or Python); (b) Can fluently consult resources such as StackOverflow and Github Issues to seek assistance and identify specialized methods and tools as needed; (c) Can work at the Advanced level in the other core programming language	(a) Works fluently with general methods and tools in at least one core programming language; (b) Comfortable using some set of specialized methods and tools with minimal support; (c) Can fluently consult resources to seek assistance and identify specialized methods and tools as needed; (c) Can work at the Intermediate level in the other core programming language	(a) Comfortable using general methods and tools with minimal support in at least one core programming language; (b) Can use some set of specialized methods and tools with the support of assistance or references; (c) Can consult resources to seek assistance and identify specialized methods and tools as needed; (d) Can work at the Introduced level with the other core programming language	(a) Can use general methods and tools with support in at least one core programming language
Communication	(a) Can design effective web-or print-ready visualization around concise caption; (b) Can verbally deliver concise communication centered on a data-driven visualization; (c) Can readily produce variable variations of the same visualization that convey the same message in different ways	(a) Can identify and visualize data in R or Python; (b) can implement effective visualization techniques to convey a concise message to a select target audience; (c) can provide effective peer feedback	Can outline an effective visualization and develop a concise written/verbal message that targets a select audience	Can identify critical features of target audience and select an appropriate communication form
Design	(a) Works in team setting to develop a data-driven solution to a real world problem; (b) Can incorporate feedback from principal stakeholder to improve final product; (c) Can effectively document the use and underlying assumptions of final product in a concise *User Guide*	(a) Can implement beta version of a data-driven solution that incorporates principal design requirements; (b) Can communicate in writing the requirements for taking beta version into a viable product	(a) Can identify ethical and sustainable considerations that differentiate a set of possible solutions; (b) Can outline and compare the principal design constraints that differentiate a set of possible solutions;	(a) Can identify critical features of a select problem that require strategic design principles; (b) Can effectively brainstorm in a team environment
Team	Can lead a team-oriented project using GitHub to share all	(a) Can set up new GIT project that is accessible across a team; (b) Uses	(a) Understands responsibilities of team leader and can implement version	(a) Can produce and share documents using available web resources

	materials internally and externally to team	Git/Github tools for sharing, documentation and version control of team data and code	control and manage team edits to increase team productivity; (b) Can effectively use web tools (Slack or MS Teams) to communicate internally and externally; (c) Can share code and data across a team	
Ethics	Can engage with background theoretical work in relevant fields (e.g., computer ethics) to articulate new concepts that characterize novel ELSI issues in data science work and its social context	Can apply and disclose key ELSI concepts embedded or applied to the broader context of data science work, including data collection, problem setting, and solution implementation	Can apply key ELSI concepts to their own work as data scientists using casuistic reasoning (analogies), can identify embedded values and norms in their work represented by key ELSI concepts	Can define or illustrate key ELSI concepts using standard case studies or examples
Applications	Can apply key ELSI concepts to develop ethically appropriate weighting or reconciliation strategies to address power dynamics, conflicting interests, or assumptions that prevent synthesis	Working with multidisciplinary and non-academic collaborators, can (a) synthesize insights and approaches and (b) identify power dynamics, conflicting interests, or assumptions that prevent synthesis	For a given problem or project, can identify disciplines and non-academic stakeholder groups with relevant expertise	Can use concepts such as wicked problems, path dependence, siloing, and standpoint theory to explain the limitations of disciplinary approaches

Appendix C – Syllabus for Capstone Project

MIST 2XX – MDSA Integrative Capstone Project

Instructor: TBD
Credit: 4 Units

This is a project-based capstone course required of all students in the Master of Data Science and Analytics program. In this course, students work in teams of 3 to 5 to solve a real-world problem through integration of the depth of knowledge from their core coursework, with the breadth of knowledge from the theme coursework. The capstone projects require teams to develop a solution that meets an industry, market, or social need through the use of data science and analytics techniques.

There are 3 basic capstone project models:

- **University Project:** Students work directly with a UC Merced faculty member, typically on a project related to problems and methods from the faculty member's lab. These projects may also involve collaboration across the campus or across the UC system. The faculty advisor is the primary project advisor.
- **Industry-University Collaboration:** These projects are typically proposed and led by UC Merced faculty in collaboration with an `industry partner. The faculty advisor is the primary project advisor, and the industry advisor is the secondary advisor.
- **Industry-Led Projects:** The Industry advisor sets the scope of work and direction, and the company supplies all resources needed for the project. The industry advisor would be the primary project lead and paired with a faculty advisor who would sign off on the project and assign the grade at its completion.

Course Format: There is no formal classroom instruction. The main activity is open-ended project work under direct supervision of the team's faculty and/or industry advisors. Separate sections of the course are created for each faculty advisor, and students enroll in their advisor's sections. Overall coordination of the course is the responsibility of the MDSA Program Director.

Program Learning Outcomes: As the capstone activity, this course aligns with all six program learning outcomes of the MDSA program.

1. **Methods**. Students will integrate appropriate quantitative, statistical, analytical, algorithmic and coding paradigms to identify knowledge management, planning and strategic decision-making solutions in different organizational and socio-technical situations. The program covers two core programming language (R and Python).
2. **Communication**. Students will selectively draw on different modes of communication (verbal, oral, graphical, code) to inform, engage and inspire in a clear and concise manner to diverse audiences comprised of experts and non-experts.
3. **Design**. Faced with a problem-oriented case, students will design resourceful and ethical informatics-based solutions by integrating data, methods and web-based technologies to strategically organize, manage, communicate, and deliver information-based services.
4. **Team**. Faced with a problem-oriented capstone project, students will gain experience integrating knowledge, skills, theory, and methods by leveraging team- and data-oriented

productivity solutions for sharing and integrating effort and data under time and other resource constraints.

5. **Ethics**. Students will understand the imperatives underlying research and data ethics, and will scope out the societal context and implications of their work by applying conceptual frameworks from the humanities and social sciences – to identify ethical, legal, and social issues surrounding data collection and analysis; to creatively develop and evaluate responses to these issues; and to implement appropriate responses.
6. **Applications**. Students will apply informatics theory and data-management methods to address boundary-spanning problems pertaining to business, management, economics, sociology, psychology, cognitive science, environmental science and engineering.

Course Learning Outcomes: By the end of the Capstone experience, students will be able to

- apply and integrate knowledge and skills of data science and analytics in real world problems **(PLO 1, 3, 6)**.
- develop a data-based solution to a real-world problem, sensitive the needs and constraints of all stakeholders **(PLO 1, 3, 5)**.
- coordinate and work with others to accomplish large projects related to data science and analytics **(PLO 3, 4, 5)**.
- communicate effectively project outcomes to stakeholders, including the potential value of a solution **(PLO 2, 5, 6)**.

Course Deliverables: There are three required deliverables:

- **Capstone Presentation**. Each team presents the results of their capstone project in 10-minute presentation followed by a 10-minute Q&A session.
- **Capstone Video**. Each team creates a 2-minute video that highlights the work.
- **Final Project Report**. Each student submits an individual final report to the faculty advisor summarizing the team project and personal contributions to it.

Grading: Faculty advisors are responsible for assigning grades, with separate grades for each individual on the project team. There are no formal rubrics for assigning the letter grade, but advisors are expected to assign a grade that reflects the overall quality and quantity of work that the student performs on the project. To some extent, that can be partially determined from the Final Project Report, but the assessment should also consider:

- Quality and timeliness of the final deliverables.
- Extent of contribution of the student to the overall project outcome.
- Consistency of effort throughout the project.

Texts: Refer to the texts from each of the core courses you have already taken. You are assumed to be knowledgeable and have access to those materials. Additional readings may be made available online and through the UC Merced Library.

Cheating and Academic Honesty: Don't cheat. Like all universities, UC Merced has a formal policy on this: http://studentlife.ucmerced.edu/what-we-do/student-judicial-affairs/academicy-honesty-policy .

Disability Services: UC Merced is committed to ensuring equal academic opportunities and inclusion for students with disabilities (see http://disability.ucmerced.edu/). If you need any assistance, please contact the instructor.

Commitment to Equity, Diversity, and Inclusion

UC Merced upholds the tenets of the University of California Diversity Statement. Equity, diversity, and inclusion arise from the mission, professional values, ethics, and commitments of our faculty, students, and staff. In particular, students have the support of their peers, faculty, and advising staff to recognize the importance of inclusion and upholding dignity, promote differences in perspective, and celebrate the identities of everyone. Campus resources are available at the Graduate Resource Center page, including support for Basic Needs, Diversity and Inclusion, Health and Wellness and Grad Family Resources.

Appendix D – Sample Course Syllabi

COGS 2XX: Methods of Data Science I
Meeting Times: 3 hours/week
Credit: 4 Units

Description: Introduces the data analytics pipeline relevant to both academic and industrial work: obtaining raw unstructured data; cleaning, organizing, merging and identifying potential pitfalls in the data; exploring and visualizing the underlying statistics; managing data for publication and reproducibility. Introduces best-practices for handling and analyzing large multi-scale datasets using examples drawn from open-data repositories. Each instance of the course will include at least one detailed case study examining the integrated ethical and technical aspects of data science work.

Course Learning Objectives: By the end of the course, students will be able to

1. Apply concepts from software engineering and philosophy of science to methodological decisions in data science.
2. Use exploratory data analysis techniques and tools to identify potential data errors and potential phenomena for further analysis.
3. Clean raw data and produce a reproducible codebook for both downstream analysis and public release of data according to FAIR standards.
4. Manage data, analysis, and outputs for reproducibility using best practices of data handling, a clear directory structure, self-documenting code, version control, and build automation.
5. Apply philosophical and data science concepts to integrated ethical-technical analysis to case studies in algorithmic injustice.

Relationship to Program Learning Outcomes and Requirements: This course provides Introduced-level alignment with at least four of the MDSA program learning outcomes, with an emphasis on 1. Skills and Methods.

1. *Skills and Methods:* The course introduces students to fundamentals of everyday data science practice (LO 1-4), including relevant software engineering concepts (e.g., object-oriented and procedural programming) and tools and techniques for cleaning, exploring, and managing data.
2. *Communication:* Data documentation and reproducibility (LO 3 and 4) necessarily have a communicative aspect: communicating to downstream users how datasets were gathered, how variables should be interpreted, and what key methodological choices were made in data gathering, data preparation, and analysis. Students become familiar with industry-standard checklists and techniques, such as the FAIR principles for public data, that facilitate this communication.
3. *Research ethics and societal context:* The course incorporates at least one *ethical-technical case study*. The case study focuses on a high-profile controversial case of applied data science. In this course, examination of the case study focuses on technical aspects of the controversial case, such as different measures of "fairness," but relates them to broader ethical, legal, and social issues. The case study may be coordinated with the Data Ethics course, so that students spend a significant amount of time in the same semester examining both technical and ethical aspects of the case. By integrating ethical and technical issues in this way, the curriculum aims to avoid the attitude that ethics is "someone else's problem." See below for an example case study.

4. *Cross-Boundary Applications:* Students conduct weekly labs using example datasets sourced from a wide variety of academic disciplines and problem contexts. Common example datasets might include house prices, public school student demographics, published data from academic research in fields such as psychology or cognitive science, dog license application records, or research funding data from the National Science Foundation. By working with this wide variety of data, students come to appreciate both the broad scope of data science and the need for domain-specific understanding.

Example Ethical-Technical Case Study: COMPAS

The case study will take 3-4 weeks, in the third quarter of the term after students have acquired the necessary technical skills.

Learning outcomes for the COMPAS case study:

- *Understand the background on the COMPAS case.*
  COMPAS is an algorithm for predicting recidivism (whether someone arrested for a crime will be arrested again in the future). It received national attention when ProPublica published an analysis of data from Broward County, FL, indicating racial bias in the algorithm's error rates. ProPublica's dataset is publicly available, enabling students to reproduce their headline findings.
- *Reproduce several fairness statistics on COMPAS data.*
  Numerous statistical measures of "fairness" have been proposed in the machine learning and algorithmic bias literatures. Given ProPublica's data (which include COMPAS predictions and recidivism "ground truth"), it is straightforward to calculate the most prominent fairness statistics.
- *Explain tradeoffs between false negatives and false positives in terms of inductive risk.*
  Inductive risk is a concept from philosophy of science, referring to the unavoidable risk of error when making an inductive inference (e.g., taking COMPAS predictions from a training set to actual application). Philosophers of science argue that the way we balance different kinds of risk of error (e.g., false positive vs. false negative predictions) should depend on the values at stake.
- *Explain the Kleinberg et al. (2016) impossibility result (not necessarily the proof) and show that it applies to the COMPAS data.*
  Critics and defenders of COMPAS have used different fairness statistics to claim that the algorithm is racial biased or unbiased, respectively. Kleinberg et al. (2016) showed that these fairness criteria could all be satisfied only if the "base rates" are the same across groups. Because the "base rate" of recidivism differs between white and black subjects in the COMPAS data, it is impossible to design an algorithm that would satisfy both critics and defenders of COMPAS. Showing this result applies to the COMPAS case is a matter of calculating the relevant fairness statistics.
- *Critically discuss profiling and thresholds as potential explanations for differences in "base rates" of recidivism.*
  Recidivism is not an intrinsic attribute of a person, but instead depends on how an individual is treated by their social environment. Who is arrested depends on who is subject to police surveillance. Insofar as black people are subject to racial profiling and lower "threshold effects," bringing them under greater police scrutiny, they are more likely to be arrested. So racial bias in policing practices, rather than racialized differences in some intrinsic "criminality," might explain differences in racial "base rates" of recidivism.

- *Critically locate discussions of fairness and recidivism in a broader social context, and identify implications for working data scientists.*
  Students might take up questions such as the following:
  - Statistical measures of fairness only consider algorithmic inputs and outputs. Should the social context that produce those "inputs" and act on those "outputs" also be taken into account? If so, how?
  - What values should ML systems designers use to balance inductive risks? Who should play a role in answering that question?
  - What responsibilities do data scientists have for who uses the data collection and prediction systems that they design, and how those systems are used?
  - Race is multidimensional and socially constructed. What implications does this have for the way racial identity data are collected, managed, and analyzed?

Assessment:

Lab assignments	25% (10x 2.5% each)
Code reviews	25% (2x 12.5% each)
Case study paper	25%
Final project	25%

- *Lab assignments:* Graded automatically as complete/incomplete
- *Code review:* Students work in pairs to review and refine their code, with a 2-page writeup reporting strengths and areas for improvement
- *Case study paper:* 5-page paper on the implications of the case study for working data scientists
- *Final project:* Working in groups of 3, students find a data source, prepare a reproducible cleaning script and exploratory data analysis, and deliver a data dictionary and 2-page report

Commitment to Equity, Diversity, and Inclusion

UC Merced upholds the tenets of the University of California Diversity Statement. Equity, diversity, and inclusion arise from the mission, professional values, ethics, and commitments of our faculty, students, and staff. In particular, students have the support of their peers, faculty, and advising staff to recognize the importance of inclusion and upholding dignity, promote differences in perspective, and celebrate the identities of everyone. Campus resources are available at the Graduate Resource Center page, including support for Basic Needs, Diversity and Inclusion, Health and Wellness and Grad Family Resources.

COGS 210(?): Statistics for Cognitive Scientists
SPRING SEMESTER 2021

INSTRUCTOR: DR. RACHEL RYSKIN | rryskin@ucmerced.edu | Office hours: XX

Course Information

This course will introduce the statistical toolkit used by researchers in cognitive science to learn from data. The focus will be on (generalized) linear models (linear regression, logistic regression, multilevel models, etc.) from both frequentist and Bayesian perspectives. Students will learn to apply these techniques to real datasets using the R programming language, create publication-quality data visualizations, and build reproducible statistical workflows.

Course Organization

Grades will be based on:

1. 6 problem sets (each worth 10 points)
2. Class participation (10 points)
3. A final paper (worth 30 points)

For the final paper, students will select a dataset and research question (from a set of options made available on the course website or their own data), visualize the data, analyze it, and write a conference-proceedings length (6 pages) paper reporting the results and conclusions. The paper will be accompanied by code for reproducing all analyses.

Course Learning Outcomes

In this course, students will learn to 1) analyze and make inferences about data, 2) evaluate the appropriateness of common statistical methods being used in the literature, 3) communicate statistical results, and 4) contribute to a reproducible research enterprise. This course contributes to the Cognitive and Information Science program learning

outcomes by promoting the 1) understanding of foundational concepts in cognitive and information sciences, 2) the skillful use of foundational methods in cognitive and information sciences and developing 3) scientific communication skills.

Resources

Course materials will be available at the course website: LINK

Books:

- R4DS: R for Data Science by Garrett Grolemund & Hadley Wickham
- LSR: Learning Statistics with R by Danielle Navarro
- SR: Statistical Rethinking by Richard McElreath
- ISL: Introduction to Statistical Learning by Gareth James, Daniela Witten, Trevor Hastie, Rob Tibshirani
- ARM: Data Analysis Using Regression and Multilevel/Hierarchical Models by Andrew Gelman & Jennifer Hill

Schedule

Week	Dates	Topic	Reading	HW
Week 1		Why Learn Statistics? / Introduction to R	LSR ch.1, 3, & 4	
Week 2		Descriptive statistics & Data visualization	R4DS ch. 3 & ch. 5	Pset 1
Week 3		Comparing groups & Null hypothesis significance testing	LSR ch. 9, 10, 11	
Week 4		Bayesian belief updating	SR ch. 1-4	Pset 2
Week 5		Linear Models	LSR ch. 15	
Week 6		Interactions	SR ch. 8	Pset 3

Week 7		Categorical predictors and contrast coding	Schad et al. (2019)	
Week 8		Power & Simulation	ARM ch. 7-8	Pset 4
Week 9		GLMs	ISL ch. 3, SR ch. 10-11.1	
Week 10		Regularization, Cross-validation, & Model Comparison	ISL ch. 5-6, SR ch. 7	Pset 5
Week 11		Multilevel Models	ARM ch. 11-13	
Week 12		*Spring Break*		
Week 13		Multilevel Models	SR ch. 13	Paper Proposal
Week 14		Multilevel GLMs	SR. ch. 14	Pset 6
Week 15		Multilevel simulation	DeBruine et al. (2020)	
Week 16		Autoregressive GLMs	Cho et al. (2018)	Final Paper

Course Syllabus
COGS 269
Global Good Studio
SPRING 2020
ONLINE: Course Website on Canvas: https://catcourses.ucmerced.edu/

Dr. Lace Padilla

Professor: **Dr. Lace Padilla**
Preferred pronoun: She/Her/Hers
Preferred addressment: Dr. Padilla
Preferred contact: Lace.padilla@ucmerced.edu
Office Hours: By appointment
Office: Online
Lab: Space.UCMerced.edu

Course Overview

Global Good Studio is a project-based graduate-level course with the goal of using data science to enact positive change in the world. At the beginning of the course, students will work in groups to identify issues impacting the world and will work backward to develop skills and knowledge to move the needle on the problems they've identified. At the end of the course, students will have produced 1-3 portfolio level projects (depending on the scope of the projects) and a communication plan for disseminating their work. Projects may include collaborations with industry and agency partners.

Skill development may include:

- data analytics and analysis
- web scraping
- data visualization
- web development
- community outreach and organization
- media development
- open science practices
- grant writing
- HCI research methods

Course Objectives and Learning Outcomes

1) Students will learn how to develop and apply their training in data science to address large-scale problems.

2) Students will develop critical discourse skills needed to be an engaged scientific community member.

3) Students will develop approaches for communicating science to the global population.

4) Students will learn modern project development practices utilized in industry.

Required Materials

Readings will be provided on CatCourses.

Tentative Timeline

Date	Topic/Description	Assignment
Week 1	Introduction to Information Visualization	
Week 2	Data Abstraction	Group development
Week 3	Fundamental Graphs and Data	Project 1 timeline
Week 4	Graphical Components and Mapping	Project 1 action items
Week 5	Applied Perception for Information	Project 1 action items
Week 6	Effectiveness of Visual Channels	Project 1 action items
Week 7	Color Perception and Color Spaces	Project 1 action items
Week 8	Using Color in Visualization	Project 1
Week 9	Applied Perception	Group development
Week 10	Visualization and Memory	Project 2 action items
Week 11	Visualization decision-making	Project 2 action items
Week 12	Cognitive Theory	Project 2 action items
Week 13	Visualization Evaluation	Project 2 action items
Week 14	Uncertainty Visualization	Project 2 action items
Week 15	Data Bias	Project 2 action items
Week 16	Project presentations	Project 2

Grading

Your grade will be made up of:

- Project 1: 25%
- Project 2: 25%
- Class Participation: 50%

Projects: Each project will be evaluated based on the goals defined in the student's project outline. Students will be graded on how effectively they achieve their personal

and group level goals. Each group will present their projects to the class for feedback and outline a communication plan for sharing their findings with the world.

Class Participation: Each week, students will self-identify their action items. These action items constitute the needed steps to complete the projects as defined by the student project timeline. Students will be graded on if they completed their action items for the week.

Other information:

If you have any questions, comments, or concerns about this course please fill out the anonymous form below. We greatly value your opinion. We will work to improve the course based on your feedback.

https://ucmerced.az1.qualtrics.com/jfe/form/SV_5w4hkW6ip6DPbTf (Links to an external site.).

COGS 2XX: Ethics & Technology

Meeting Times: 3 hours/week

Credit: 4 Units

Description: Introduces students to three kinds of work in data ethics: ethical analysis of existing and proposed data science products or processes, the ethical theories that justify and facilitate moral critiques, and the production of ethical evaluations of data science products and processes. Core concepts covered include privacy, big data collection & analysis, algorithmic bias, data ownership, and the social responsibilities of engineers.

Course Learning Objectives: By the end of the course, students will be able to

1. Apply concepts from the fields of normative and applied ethics to the design and implementation of data science products and processes, and other computing technologies.
2. Use philosophical analysis and structured normative evaluation tools (i.e., matrices) to identify potential areas of moral concern.
3. Produce comprehensive analysis and corresponding evaluation of new and existing data science processes and products (e.g., expert learning algorithms) to meaningfully consider how a range of vested stakeholders are likely to be affected if these are implemented.
4. Be introduced to and evaluate existing best practices documents about research ethics, data collection, and have the opportunity to improve on these through class work.
5. Apply philosophical and data science concepts to integrated ethical-technical analysis to case studies in near term artificial intelligence.

Relationship to Program Learning Outcomes and Requirements: This course provides Introduced-level alignment with at least four of the MDSA program learning outcomes, with an emphasis on Ethics.

1. *Methods:* The course introduces students to philosophical analysis, with a focus on teaching students how these methods can disclose embedded values, norms, and assumptions in the products and processes of data science (LO 1, 2). This is a crucial deliverable with respect to the development of socially responsible decision-making solutions.
2. *Communication:* The course teaches students to use ethical matrix tools that enable structured analysis of new technologies and technology policies (LO 3). The matrices themselves serve as artifacts of conversations about the predictable consequences of these technologies, thus serving a dual role of being a tool for analysis and a shareable and accessible artifact for communicating areas of benefit and concern.
3. *Ethics:* The course is designed such that students are exposed to a wide range of historical and contemporary case studies from computer ethics, technology ethics, and data science (LO 5). One assignment that may occur annually involves a cross-class assignment with the Data Methods I course involving a high-profile and controversial case of applied data science, the COMPAS algorithm used to predict recidivism risk. The field of computer ethics can be traced as early as 1950, with the articulation of the coming computer revolution by Norbert Wiener. As Wiener made clear, revolutions lead to significant conceptual and policy vacuums. The methods of analytic philosophy are well-suited to address these needs, and students will gain these skills with this particular end in mind. It has become clear, for

instance, that our pre-existing theories of privacy are not well-suited to the virtual public-private spheres that we find on, for example, social media, leading to serious controversies about what is or is not appropriate to post on platforms like Twitter. Other conceptual and policy vacuums exist around informed consent for things like smart phone applications that are an important source of the big data sets at the heart of data analytics.

4. *Applications:* The field of data science ethics is introduced to students as a descendant of the field of computer ethics. Computer ethics is best pursued as an interdisciplinary research area, that can draw on the expertise of philosophers, computer scientists, and social scientists. This interdisciplinary work is important in order to surface and disclose the values and norms embedded in the development of new technologies like data analytics, and also for determining solutions to the normative and technical problems surfaced by these teams. The methods of analysis used in the course (LO 1, 2, and 3) rely on drawing from this range of expertise in order to provide informed, insightful, and ideally realisable solutions to the complex problems of data science in practice.

Example Course Reading Schedule

Please note the following schedule divides each meeting into two parts, with different readings assigned for the activities and conversation in each half.

The total number of pages each week is roughly 60 pages. This is low by the standards of a typical philosophy graduate seminar, but with the applied focus of the course it is appropriate.

Wk	Reading
1	Ward Bynum, “Ethical Challenges to Citizens of “The Automatic Age””; Wiener, “Some Moral and Technical Consequences of Automation”. Moor, “What is Computer Ethics?”; Brown, “Is There An Ethics of Computing?”.
2	Mill, Utilitarianism (selections); O’Neill, Simplified Kantian Ethics. Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor,* Ch. 1.
3	Rawls, Theory of Justice (selections); Shue, “Global Environment And International Inequality”. Gorniak-Kocikowska, “The Computer Revolution and the Problem of Global Ethics”; van den Hoven & Rooksby, “Distributive Justice & Information”.
4	Scanlon, What We Owe to Each Other (selections); Johnson, “Do Engineers have Social Responsibilities?”; Nissenbaum, “Computing and Accountability”.
5	Friedman and Nissenbaum, “Bias in Computer Systems”; Brey, “Method in Computer Ethics”. Ruha Benjamin, *Race After Technology,* Ch. 1, 2.
6	Johnson, “Rethinking The Social Responsibilities of Engineers”; Pirtle & Szajnfarber, “On Ideals For Engineering in Democratic Societies”. Read: Lipworth, Mason, & Kerridge, “Ethics and Epistemology of Big Data”, and select one of the contributions to the Symposium introduced in the piece to read & contribute to class discussion.
7	Feinberg, “The Nature and Value of Rights”; Shue, Correlative Duties. Stemplowska, Human Rights

8	Belmont Report; Menlo Report. FB Emotional Contagion study report; boyd, Untangling research and practice: What Facebook's "emotional contagion" study teaches us.
9	Nguyen, "Epistemic Bubbles and Echo Chambers"; Gunn, "Echo Chamber, Filter Bubbles, & Online Communities". Vallor, *Technology & The Virtues,* Ch. 1, 7.
10	Thomson, "The Right To Privacy"; Solove, "Misunderstandings of Privacy". Custers et.al, "Consent and Privacy"; Eubanks, *Automating Inequality,* Ch 3.
11	Propublica, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing; Kleinberg et.al. "Inherent Trade-Offs in the Fair Determination of Risk Scores". Gunn & O'Neil, "Near Term AI and the Ethical Matrix"; Eubanks, *Automating Inequality,* Ch 4.
12	[Week for cross-class assignment with Data Science Methods I]
13	[Readings for the rest of the semester to be determined with the cohort]
14	
15	
16	

Requirements

The course has four main requirements: 1. Readings, 2. Seminar participation, 3. Contributing to the annotated bibliography class project, and, 4. A research paper or project.

- Weekly readings for classes
- Attendance and participation in weekly classes
- Semester long annotated bibliography
    - The annotated bibliography has two components:
        - Weekly 70 word annotations on *at least one* of the readings
        - One in-class presentation with a 400 word commentary paragraph submitted the week after your presentation
            - *At the end of the semester, I will compile all of the annotations and commentary paragraphs into one document that will then be sent to all members of the class.*
- A term paper or similarly extended research project. Options for this assignment to be discussed, one may opt to complete an in depth ethical analysis assignment with corresponding evaluation in accordance with the Ethical Matrix tool used over the course of the semester.
    - The term paper (or similar) also has two components:
        - A mid-term draft and the final paper.

ECON 2XX: Methods of Data Science II
Meeting Times: 3 hours/week
Credit: 4 Units

Description: The course will mainly focus on two aspects. First, address the differences and complementarities between machine learning and causality estimation techniques. The benefits of classification trees or random forests will be compared to the ones of causal estimations like instrumental variable regressions. Second, text mining techniques like neural networks will be introduced in the context of real open-access datasets. The social and ethical implications of predictive analytics methods will be discussed.

Course Learning Objectives: By the end of the course, students will be able to

6. Apply a large of estimators for prediction analysis in applied datasets knowing the optimization algorithms
7. Understand the conceptual difference between estimator aiming at prediction versus causality
8. Work with large datasets and use data mining techniques
9. Express in computer language the resolution of complex text and prediction problems
10. Get introduced to the idea of forecasting with text data

Relationship to Program Learning Outcomes and Requirements:

This course provides alignment with at least four of the MDSA program learning outcomes:

1. *Methods*: The course will introduce students to the fundamentals of prediction algorithms, both theoretically and in an applied manner. Further, it will heavily rely on Python to build knowledge on how to optimally use algorithms to analyze text data.

2. *Communication:* As part of the tasks will require teamwork, the course should also enhance communication skills. One key goal is to build interdisciplinary study groups such that students have to admit their communication tools to the target audience.

3. *Design*: Students will have to design solutions to data problems and evaluate which approach is more suitable for each dataset and research question. Given the variety of statistical available tools, the reasoning for why a given statistical approach is chosen becomes a key element of the course.

4. *Applications*: Students will apply the theoretical concepts discussed in class in several data-driven exercises spanning different academic disciplines. Common example datasets might include newspaper articles, patents data, or financial data, among others.

Course Schedule:

PART I: Machine Learning and Causality

We will follow the book by Gabor Bekes and Gabor Kezdi “Data Analysis for Business, Economics, and Policy”:

Chapter 13: A Framework for Prediction

Chapter 14: Model Building for Prediction

Chapter 15: Regression Trees

Chapter 16: Random Forest and Boosting

Chapter 17: Probability Prediction and Classification

Further Topics to be discussed:

Lucas Critique in the context of Big Data applications

Conceptual differences between Machine Learning and Causal Estimation

Applications of Machine Learning in Social Sciences Research

PART II: Text Analysis

We will use Python to learn how to analyze text data:

Syntax and Basic Data Structures

Word-counting Approaches

Neural Networks for Text Analysis (Latent Dirichlet Allocation; word-to-vec)

Applications of Text Analysis for Research in Political Science and in Innovation

Nowcasting and Forecasting Economic and Financial Data

Appendix E – Independent Assessment: Demand, Jobs, Competition

MAIN REPORT

UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

MASTER'S OF DATA SCIENCE AND ANALYTICS MARKET ANALYSIS REPORT

Executive Summary

The University of California, Merced has proposed a one-year Master's of Data Science & Analytics program that successful students can apply in diverse information-driven public and private industry settings. The program includes education in practical economic principles, principles and best practices for the ethical use of data, effective communication with data, introduction to scripting languages for data management and visualization, design of interactive web- and smart-phone-based platforms to facilitate organizational systems thinking and real-time decision-making, and best practices and management tools for data-intensive team-work.

The UC San Diego Extension Center for Research and Evaluation (CR+E) conducted a labor market analysis to quantify prospective student and workforce demand for the proposed UC Merced Master's of Data Science & Analytics program. The analysis included three distinct, but related components: 1) a competitive landscape to understand existing programs in California and across the United States that are similar to the proposed program, 2) an occupational outlook for graduates of the proposed program, and 3) a survey of prospective students to illicit their feedback and interest in the proposed program. Key takeaways from the market analysis include:

- There are numerous master's-level data science and analytics programs available in California and across the United States
- The demand for data scientists is growing rapidly with the number of job postings requiring data science skills increasing to over 1 million in 2019—more than a 600% increase since 2010
- The number of available jobs that require data science skills is seven times greater than the number of jobs with data science titles
- California has a greater demand for professionals with data science skills than any other state in the union
- Of the 300 prospective students surveyed, 56% said the proposed UC Merced program is of interest to them, and another 39% said the program might be appealing to them; only 5% of respondents said they are not interested in the program
- Survey respondents whose highest degree is in engineering expressed the most definitive interest (62%) in the proposed program, followed by those with a highest degree in the social sciences (57%)
- Survey respondents between the ages of 21 and 28 years expressed the most interest in the proposed data science program
- Gaining knowledge of data science, increased job security, and professional advancement were the most frequently cited reasons for proposed program appeal among prospective students
- Skills learned, cost, and academic reputation were cited most frequently as factors that would influence enrollment decisions

- Pricing sensitivity analyses suggest the optimal price point for the proposed program is in the $50-60k range.

Competitive Landscape

The UC San Diego Extension Center for Research and Evaluation (CR+E) identified six master's programs that were similar in focus to the proposed UC Merced Master of Data Science & Analytics offering. Table 1 provides information on the programs identified, including the institution, program name, location, length, and cost. The research team also evaluated an additional six programs that shared some similarities with the proposed program, but these were ultimately not analyzed because the focus, structure, and/or content did not align closely enough with the proposed program. These six programs are listed in Appendix A.

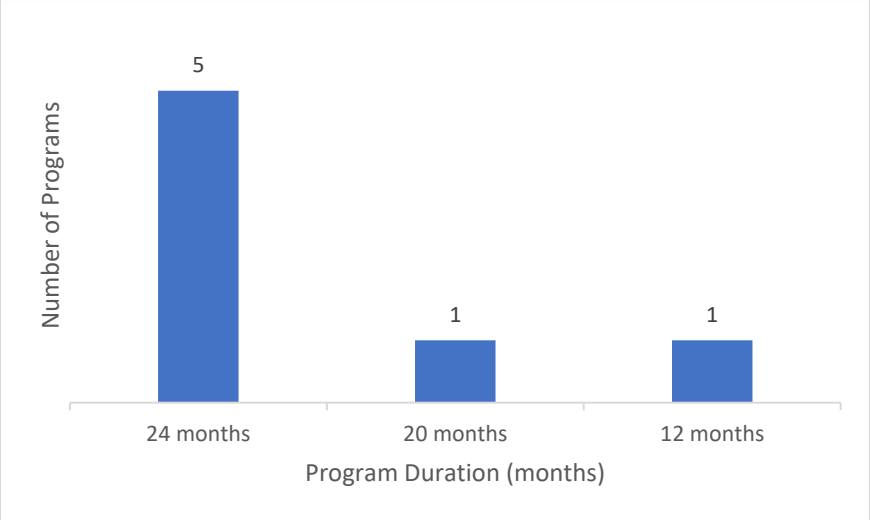
Table 1. Comparable Data Science Master's Programs at Other Institutions

Institution	Name of Program	Location	Program Length (full time)	Semester Units	In-State Tuition
UC Berkeley	Information Management and Systems (MIMS)	Berkley, CA	20	48	$59,552
UC Santa Barbara	MA in Statistics (Data Science track)	Santa Barbara, CA	24	28	$27,108
Georgetown University	MS in Data Science Public Policy (MS-DSPP)	Washington, D.C.	24	39	$74,310
University of Southern California	MS in Environmental Data Science	Los Angeles, CA	20 or 24	36	$90,700
University of Michigan	MS in Environmental Informatics	Ann Arbor, MI	24	42	$93,000
UC Davis	MS Business Analytics (MSBA)	Davis, CA	12	29.3	$92,257
UC Merced	***MS Data Science & Analytics***	***Merced, CA***	***12***	***32***	***$45,000+***

Program Length

Program lengths for comparable master's programs ranged from 12 months to 2 years of full-time study. The average time to complete a master's in data science and analytics, or similar degree, is 21 months, with a mean requirement of 37-semester units of coursework. Figure 1 shows completion time for the six programs evaluated.

Figure 1. Completion Time for Comparable Master's Programs (In Months)



Internships, Theses, and Capstones

Most programs do not require students to take internships, but UC Berkeley and Georgetown highly recommend them for building work experience. Master's programs in data science tend to be project-based to provide real world experience. At UC Santa Barbara, students must pass an applied statistics qualifying examination, and complete an instructor-approved analysis project to fulfill degree requirements. A capstone is required for the proposed UC Merced program, and several other programs evaluated. Table 2 shows requirements by program.

Table 2. Comparable Program Requirements

Institution	Internship	Thesis	Capstone
UC Berkley	■		■
UC Santa Barbara			
Georgetown University	■		
University of Southern California			
University of Michigan			■
UC Davis			■
UC Merced			■

Program Format

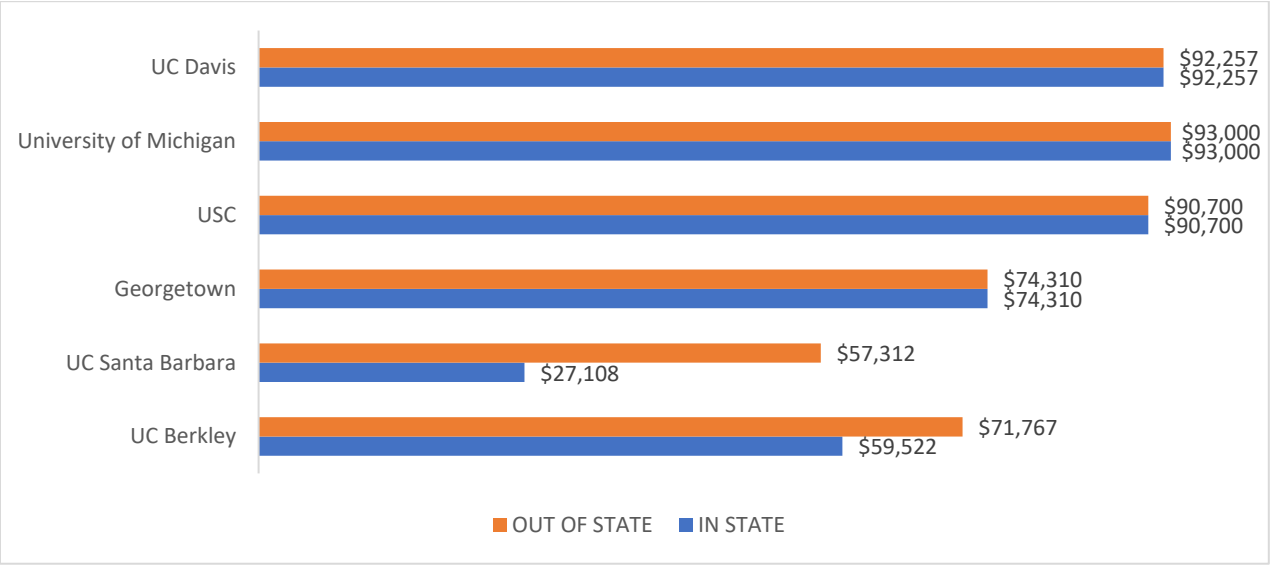
All of the programs evaluated were strictly onsite offerings, geared toward full-time study[68].

[68] The Covid-19 pandemic may influence the format of offerings

Tuition

The average, in-state tuition cost of the programs evaluated was $72,816. Only two of the six programs charged a different rate for out-of-state tuition: UC Berkley and UC Santa Barbara. Figure 2 shows tuition for each comparable program by residency status.

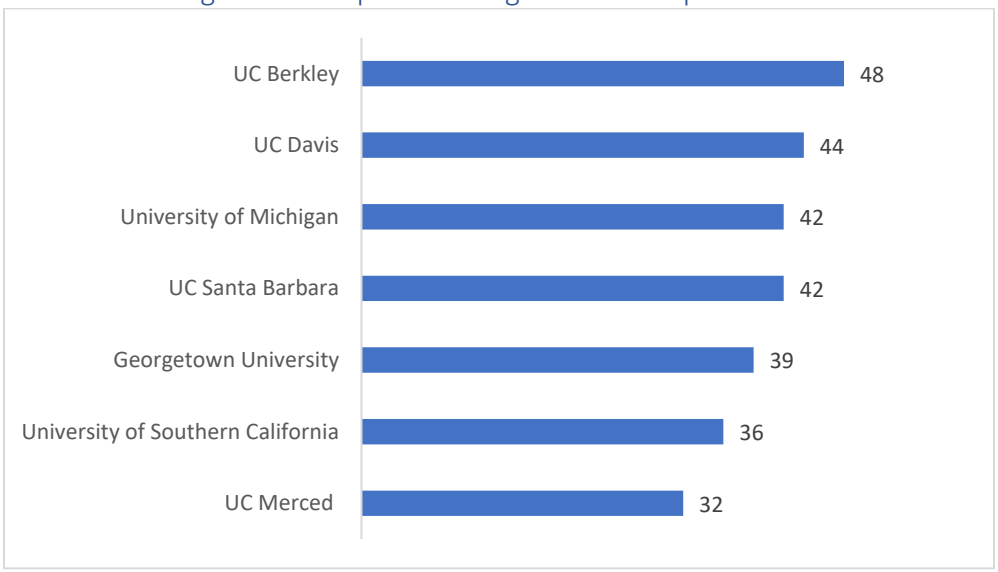
Figure 2. Comparable Program Tuition by Residency Status



Curriculum Comparison

Master's programs in data science require 37-semester units, on average, to complete. The proposed UC Merced program can be completed in 32-semester units as shown in Figure 3.

Figure 3. Comparable Program Unit Requirements



The required course worked for comparable programs is outlined in Table 3. Figure 3 is a map of the proposed curriculum as it compares with existing, comparable programs. Course titles vary across institutions, but concepts are similar.

Table 3. Required Coursework for Comparable Programs

Institution	Required Courses
UC Berkley	▪ Introduction to Data Science Programming ▪ Research Design & Application for Data and Analysis ▪ Statistics for Data Science ▪ Fundamentals of Data Engineering ▪ Applied Machine Learning ▪ Three of the following: Data Visualization, Behind the Data: Humans and Values, Experiments and Casual Inference, Deep Learning in the Cloud and at the Edge, Machine Learning at Scale, Natural Language Processing with Deep Learning, Statistical Methods for Discrete Response, Time Series, and Panel Data ▪ Capstone
UC Santa Barbara	▪ Advanced Statistical Methods A-B-C series ▪ Seminar and Projects in Statistical Consulting ▪ Statistical Data Science ▪ 2 Elective courses in Data Mining, Computational Techniques in Statistics and/or Big Data Analytics
Georgetown University	▪ Quantitative Social Sciences ▪ Foundations of Public Policy ▪ Civic Data Science ▪ Ethics and Law ▪ Communications ▪ 6 additional elective credits
USC	▪ Principles of Programming for Informatics ▪ Introduction to Computational Thinking and Data Science ▪ Overview of Data Informatics in Large Data Environments ▪ Interdisciplinary Approaches to Environmental Studies ▪ Environmental Risk Analysis ▪ Concepts for Spatial Thinking; ▪ 8 additional elective credits
University of Michigan	▪ GIS ▪ Remote Sensing; Ecology: Concepts and Applications ▪ The Science and Practice of Social Change ▪ 3 credits from the Integrated Analytics Methods & Skills courses ▪ Statistics ▪ 6 credits of elective coursework ▪ Capstone
UC Davis	▪ Introduction to Business Analytics ▪ Organizational Issues in Implementing Analytics ▪ Application Domains ▪ Data Management ▪ Data Visualization ▪ Data Design and Representation

	▪ Big Data ▪ Statistical Reasoning and Exploration ▪ Foundations of Analytics ▪ Advanced Statistics; Machine Learning ▪ Topics in Business Analytics ▪ Analytic Decision Making ▪ Topics in Business Analytics ▪ Practicum: Initiation ▪ Elaboration, Analysis and Implementation
UC Merced	▪ *Core Data Science* ▪ *Core Data Ethics* ▪ *Core Statistics & Probability* ▪ *Core Data Communications & Visualization* ▪ *Interactive Data Development* ▪ *Capstone* ▪ *Elective credits based on themes selected*

Figure 3. Curriculum Map

	UC Merced	UC Berkeley	UCSB	Georgetown University	USC	University of Michigan	UC Davis
Data Science							
Data Ethics							
Statistics & Probability							
Data Communications & Visualization							
Interactive Data Development							

Prerequisites

Prerequisites vary by program, with some requiring quantitative skills and others requiring social science or STEM related degrees. A listing of prerequisites for comparable programs is presented in Table 4.

Table 4. Prerequisites for Comparable Programs

Institution	Prerequisites
UC Berkley	▪ Undergraduate degree ▪ Successful work experience in relevant fields

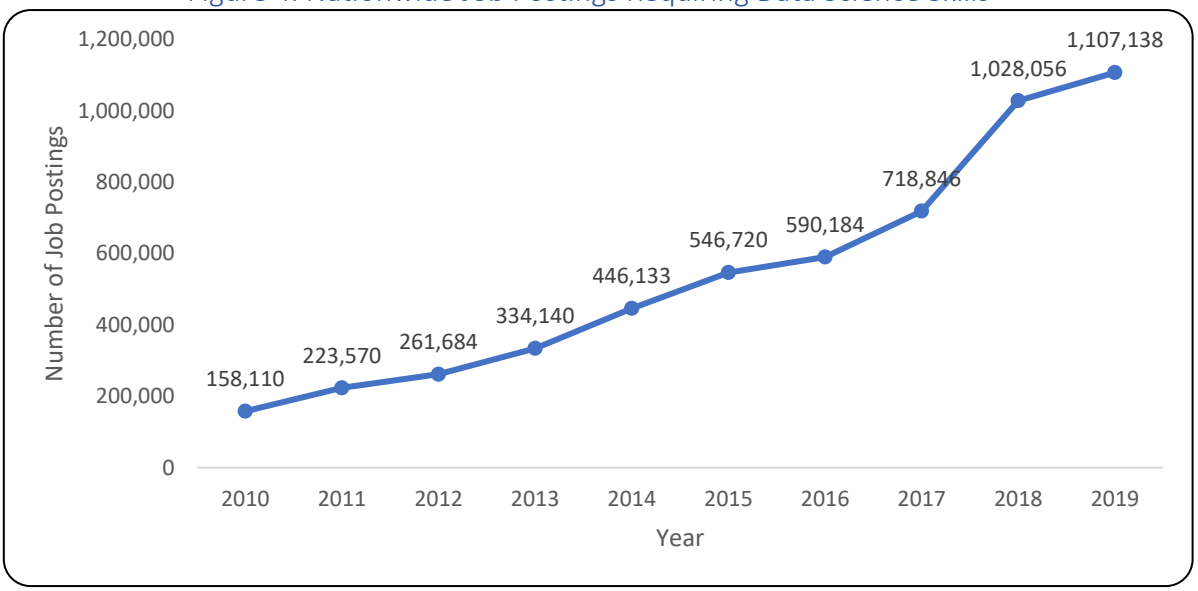
UC Santa Barbara	▪ Undergraduate degree in statistics, mathematics, or other strong quantitative field ▪ Coursework in the following areas: calculus-based probability, law of large numbers and central theorems, theory of statistical inference, hypothesis tests and regressions, linear algebra including vector spaces, eigenvalues, and eigenvectors
Georgetown University	▪ Undergraduate degree ▪ Experience in college-level calculus ▪ Demonstrated technical skills in computer science, statistics, and advanced mathematics
USC	▪ Undergraduate degree in STEM or a related social science ▪ Programming experience or a strong math background
University of Michigan	▪ Undergraduate degree from an accredited university
UC Davis	▪ Undergraduate degree with completed coursework in the three primary areas: Computing, Mathematics and Statistics and Probability ▪ Aptitude for quantitative-oriented coursework and careers
UC Merced	▪ ***Undergraduate degree*** ▪ ***Other requirements still being discussed***

Occupational Outlook

Occupational Outlook: Nationwide

The demand for skilled data scientists is growing rapidly with the number of job postings requiring data science skills increasing over 600% since 2010. Figure 4 shows jobs postings that require the following skills: *data science*, *machine learning*, *data analysis*, *tableau*, *predictive models*, *big data*, *quantitative data analysis*, and *data visualization*. According to the labor market analytics tool, Burning Glass, there were more than a million such postings in the U.S. in 2019.

Figure 4. Nationwide Job Postings Requiring Data Science Skills[69]



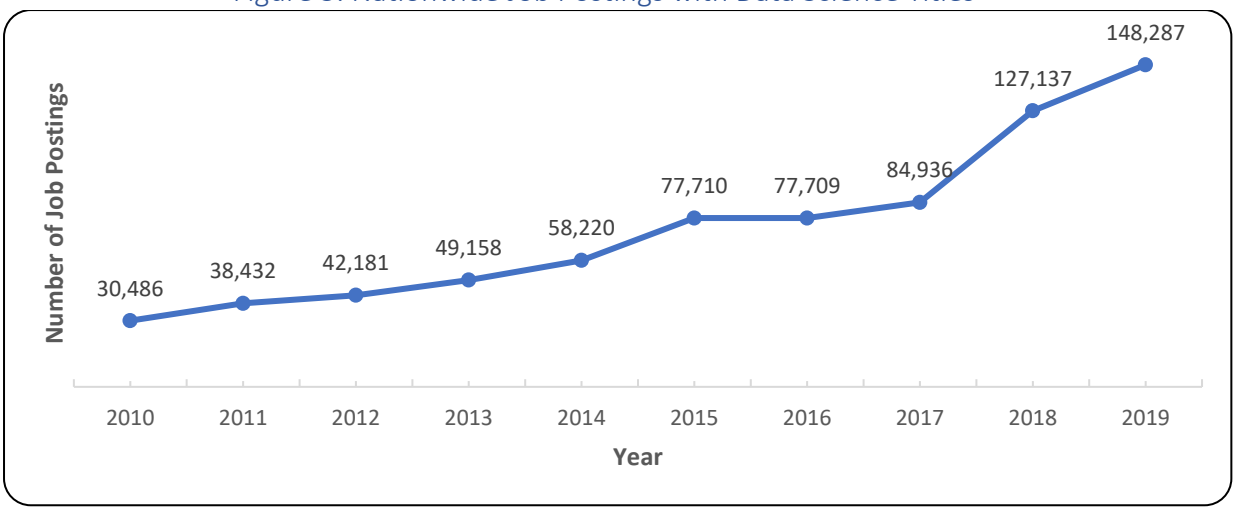
In parallel to the rapid increase in demand for data science skill sets, as shown in Figure 4, there has been a dramatic increase in the demand for jobs with data science titles. Figure 5 shows the demand for data science professionals with the following titles:

- Data Scientist
- Data Mining Analyst
- Data Engineer
- Business Management Analyst
- Database Architect
- BI Analyst
- Research Associate

In 2019, there were over 148,000 job postings for these 7 titles; nearly a 700% increase since 2010. The broad applicability of data science skill sets, as shown in Figure 4, is notable—there are far more jobs that require data science skills than those that have data science titles.

[69] Source: Burning Glass. skills include: Data Science, Quantitative Data Analysis, Machine Learning, Data Analysis, Tableau, Predictive Models, Big Data, and Data Visualization

Figure 5. Nationwide Job Postings with Data Science Titles[70]



The U.S. Bureau of Labor Statistics (BLS) uses a Standard Occupational Classification (SOC) system to organize and track occupations. Table 5 includes the SOC codes most closely associated with data science occupations. All occupations listed are projected to grow at a rate much faster than the overall average occupational growth of 5% over the 2018-2026 time span. It is notable that the salaries for these occupations are all well above the national median.

Table 5. Salary and Employment Projections by SOC Group (All Education Levels)[71]

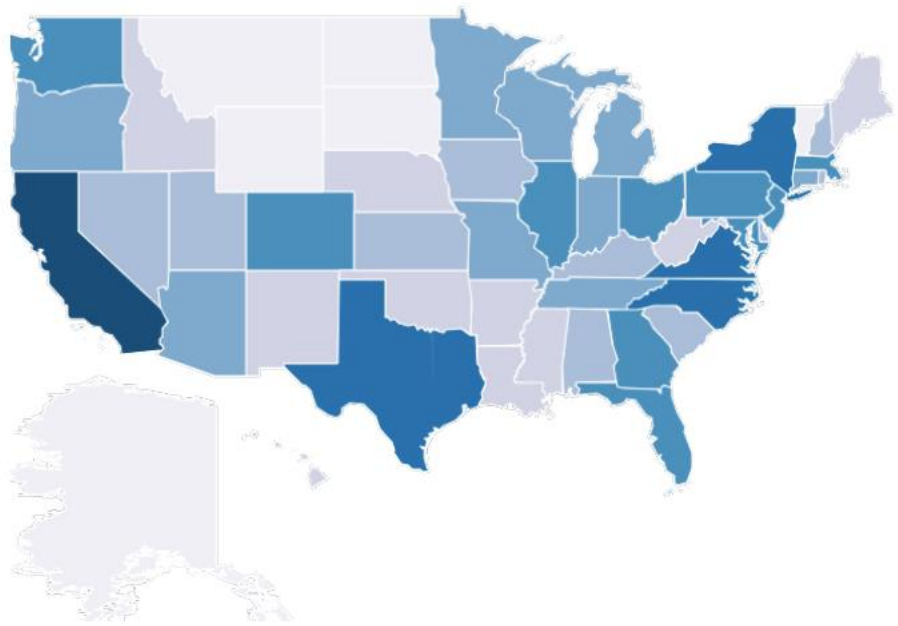
Occupational Title (SOC)	Number Employed 2018	Projected National Change in Employment, 2018-2026	Median Salary 2019
Actuaries	25,000	20%	$108,350
Economists	21,000	8%	$105, 020
Mathematicians	2,900	26%	$92,030
Statisticians	44,400	31%	$91,160
Operations Research Analysts	109,700	26%	$84,810
Data Scientists	31,700	16%	$94,280

Figure 6 is national heat map that shows the concentration of job postings that require data science skills by state. Darker shades indicate more jobs, while lighter shades indicate fewer jobs.

[70] Source: Burning Glass. Job titles include: Data Scientist, Data Mining Analyst, Data Engineer, Business Management Analyst. Database Architect, BI Analyst, or Research Associate

[71] Source: Bureau of Labor Statistics

Figure 6. National Heat Map of Jobs Postings Requiring Data Science Skills (Aug 2019-2020)[72]



Occupational Outlook: California

As was shown in Figure 6, the demand for jobs that require data science skills in California is "very high" (as defined by Burning Glass labor market analytics tool). Table 6 shows number of postings for the job titles with the most in demand data science skills in California over the past 12 months.

Table 6. California Jobs that Require Data Science Skills (Aug 2019-July 2020)[73] [74]

Job Title	Postings	% Postings
Data Analyst	10,598	5%
Software Development Engineer	9,270	4%
Business Analyst	4,493	2%
Data Engineer	3,686	2%
Data Scientist	2,602	1%
Machine Learning Engineer	2,191	1%

[72] Source: Burning Glass; Skills included: Data Science, Python, Machine Learning, Data Analysis, Big Data, Java, Tableau, and Predictive Models

[73] Source: Burning Glass; Skills included: Data Science, Big Data, Quantitative Data Analysis, Data Visualization, Machine Learning, Data Analysis, Tableau, and Predictive Models

[74] Interns were also include as jobs that require data skills with 4,493 (or 2%) of job postings. Excluded from table as these are usually temporary

Financial Analyst	1,879	1%
Data Architect	1,577	1%
Senior Data Engineer	1,531	1%
Business Consultant	1,454	1%

Table 7 shows California job postings with data science titles over the past 12 months.

Table 7. California Jobs with Data Science Titles(August 2019-July 2020)[75]

Occupational Title (SOC)	Postings	% Postings
Data Engineer	4,968	14%
Research Associate	2,951	9%
Data Scientist	2,605	8%
Senior Data Engineer	2,043	6%
Senior Data Scientist	1,173	3%

SURVEY

Background and Demographics

To measure prospective student interest and preferences regarding the proposed UC Merced Master's of Data Science & Analytics program, 301 survey responses were collected through a survey panel in August 2020. The sample was specified such that all respondents were between 18 and 39 years of age, with 78% of the sample drawn from respondents 18-28 years of age, and 21% from the 29-39 year-old age group who the UC San Diego Center for Research and Evaluation (CR+E) conceptualize as potential "career changers". Respondents were distributed across the U.S., with a focus on California (39% of sample). To qualify for the survey, respondents had to have a bachelor's degree or higher; they were screened out if they failed the attention check question. Approximately 54% of qualified respondents were female and 45% were male, less than 1% identified as non-binary or transgender. The survey began with a description of the proposed program, and then respondents were asked a series of questions about appeal of the degree, enrollment decisions, and pricing.

In addition to the responses collected by CR+E, UC Merced Alumni Relations collected 100 responses from alumni; see separate slide deck.

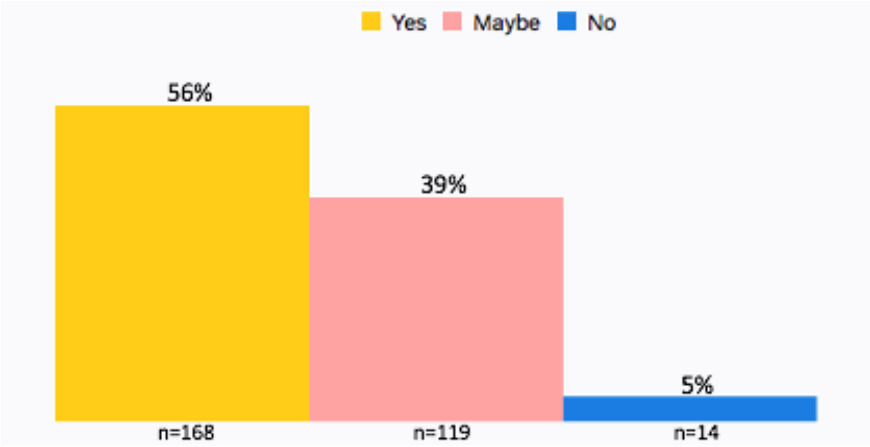
Proposed Program Appeal

More than half (56%) of the young adults surveyed expressed interest in the proposed UC Merced Master's of Data Science & Analytics program. An additional 39% said the

75 Source: Burning Glass; Job Titles Included: Data Scientist, Data Mining Analyst, Data Engineer, Business Management Analyst. Database Architect, BI Analyst, or Research Associate

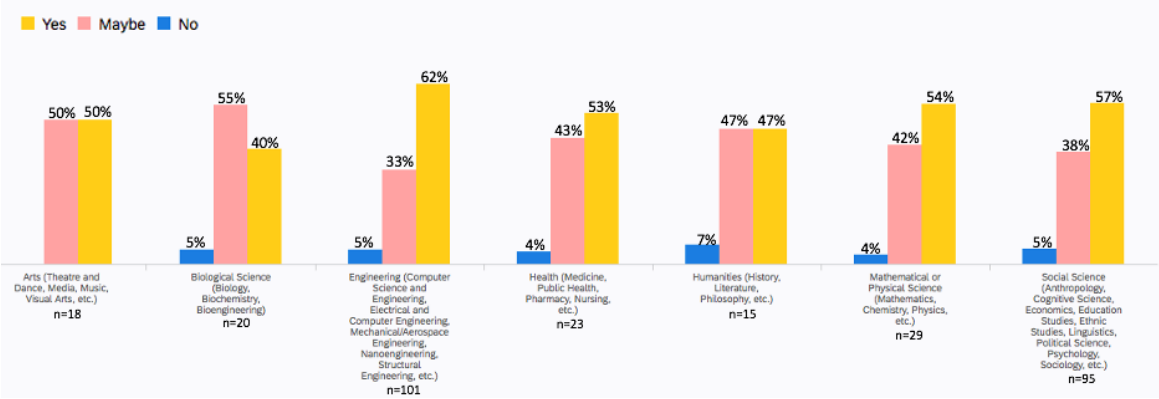
program might be appealing to them, and only 5% of respondents said the program was of no interest to them; these trends are shown in Figure 7.

Figure 7. Appeal of Master's of Data Science & Analytics Degree



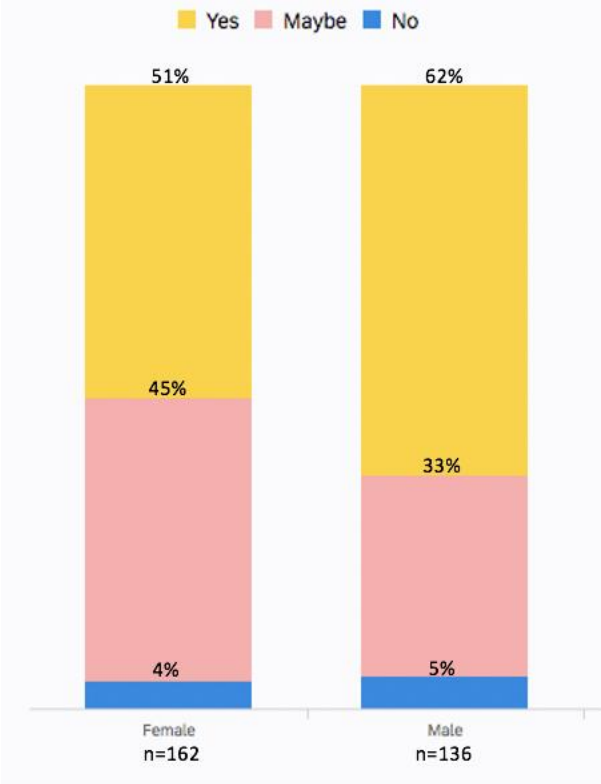
Respondents whose highest degree is in engineering expressed the most interest in the proposed program (62%), followed by those with highest degree in social science (57%). Figure 8 shows degree appeal by area of highest degree.

Figure 8. Degree Appeal by Highest Degree Area



Male respondents (62%) expressed more interest in the proposed Master's of Data Science & Analytics program than female respondents (51%), as shown in Figure 9.

Figure 9. Degree Appeal by Gender



The age group that expressed the most interest in the proposed program were those 21 – 24 years old (60%), followed by the 25 – 28 year-old age group (57%), as can be seen in Figure 10.

Figure 10. Degree Appeal by Age

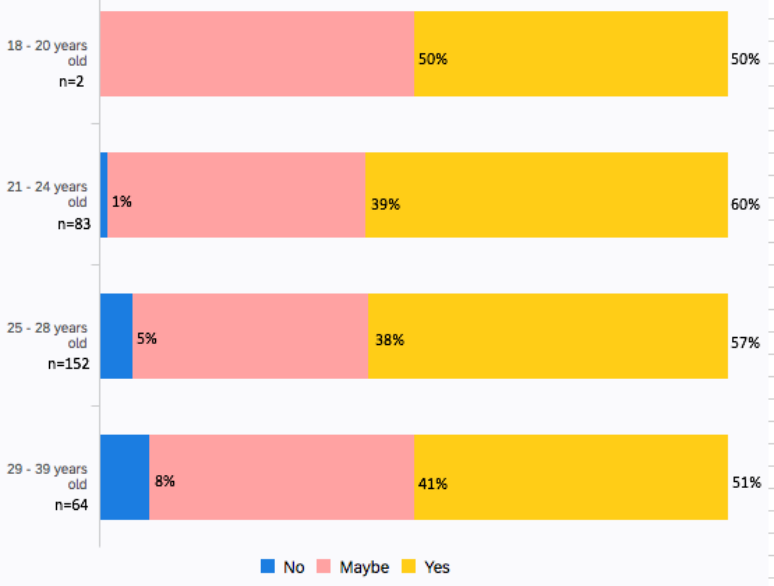
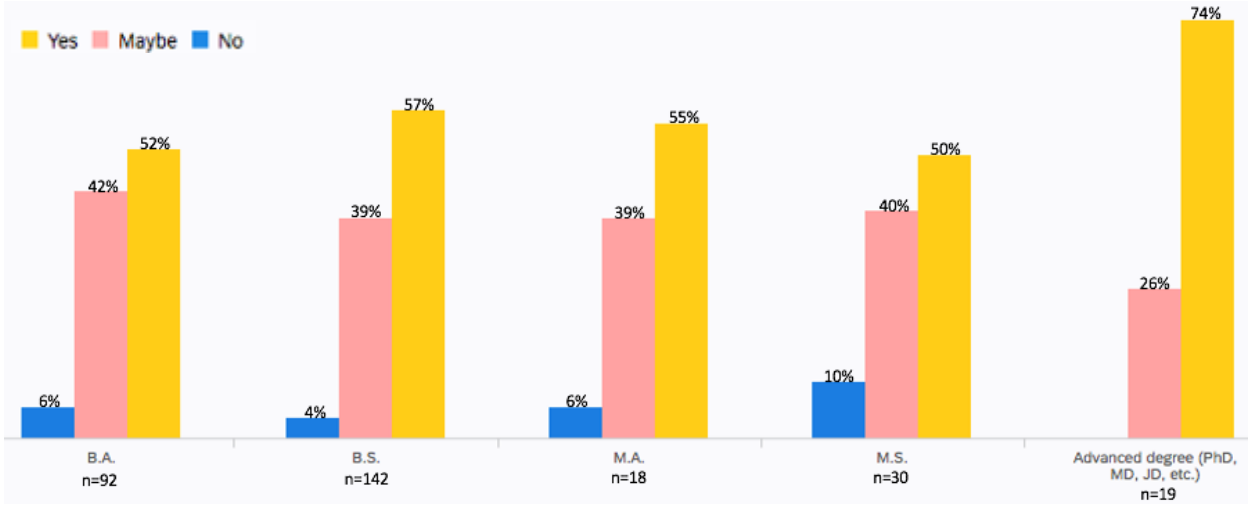


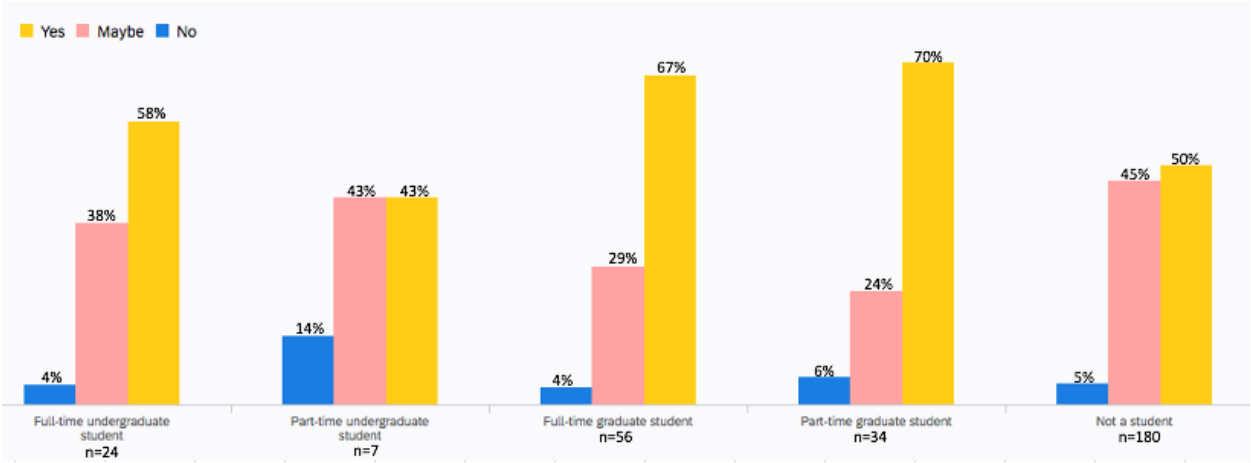
Figure 11 shows proposed degree appeal by highest degree attained. Only people with a bachelor's degree or higher were permitted to complete the survey. Respondents with bachelor's degrees expressed slightly more interest in the proposed Master's of Data Science & Analytics program than those with master's degrees. Although the sample size is small (n=19), it is notable that respondents with advanced degrees expressed a high level of interest in the proposed program.

Figure 11. Degree Appeal by Educational Attainment



In figure 12, proposed degree appeal by student status is presented. Interestingly, current graduate students expressed the most interest in the proposed program.

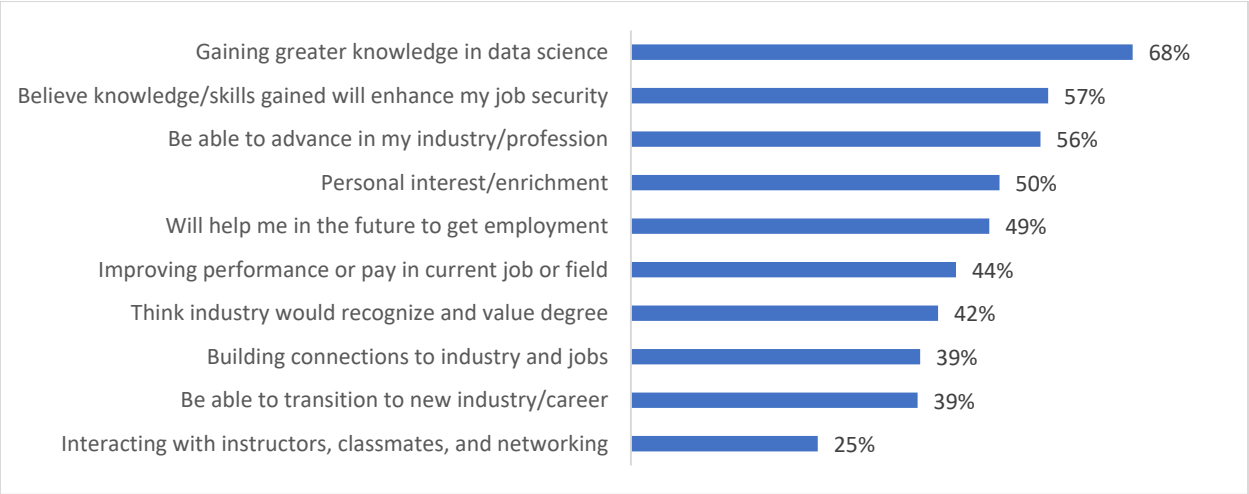
Figure 12. Degree Appeal by Student Status



Reasons for Program Appeal

Of the respondents who said they were, or might be, interested in the proposed UC Merced program, the top reason cited for program appeal was gaining greater data science knowledge (68%), followed by the belief that the skills gained through the program would enhance job security (57%), and promote career advancement (56%). Figure 13 shows these trends.

Figure 13. Reasons Proposed Program is Appealing



Among the 5% of respondents who said the proposed program was not appealing to them, the most frequently cited reason was not having time at this point in their career/life (41%), followed by cost (35%), as shown in Table 8.

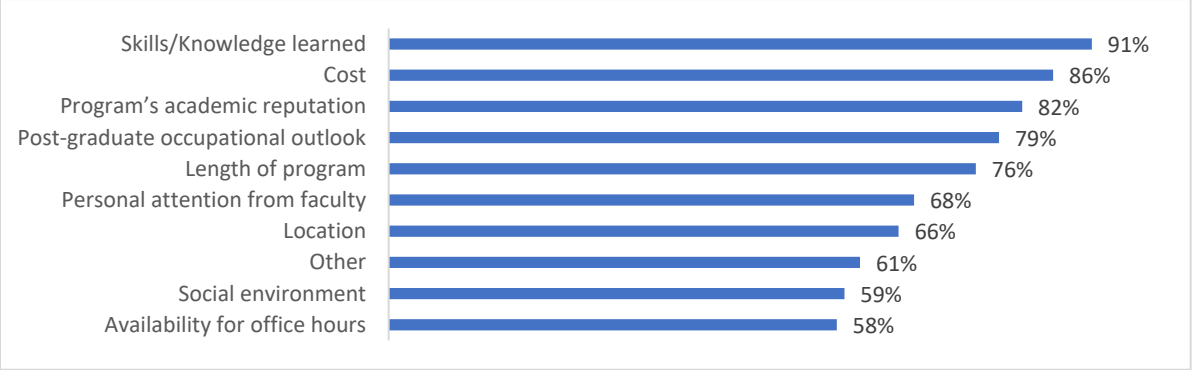
Table 8. Reasons Proposed Program is Not Appealing (n=14)

Reason	Percent of Responses
No time at this point in my career	43%
Cost	36%
Do not think will be valuable in my transition to new industry/career	29%
Do not think it will help me in the future to get employment	29%
Do not think this degree would improve my pay in current job or field	21%
Do not think it will help me advance in my industry/profession	21%
Other	21%
Do not think this degree would improve my performance	14%
Masters not essential to getting desired job	14%
Do not believe knowledge/skills gained will enhance my job security	14%
Do not think industry would recognize and/or value degree	7%

Factors Influencing Enrollment

Skills learned (91%), cost (86%), and program's academic reputation (82%) were the most cited factors that would influence enrollment in the proposed program, as shown in Figure 14.

Figure 14. Factors Influencing Enrollment in Proposed Program



Detailed descriptions of proposed core courses were provided to survey respondents, and they were asked to rate their level of interest in each course. As shown in Figure 15, high levels of interest were reported across the core courses, with *Data Communication & Visualization* and *Data Science* rated most highly.

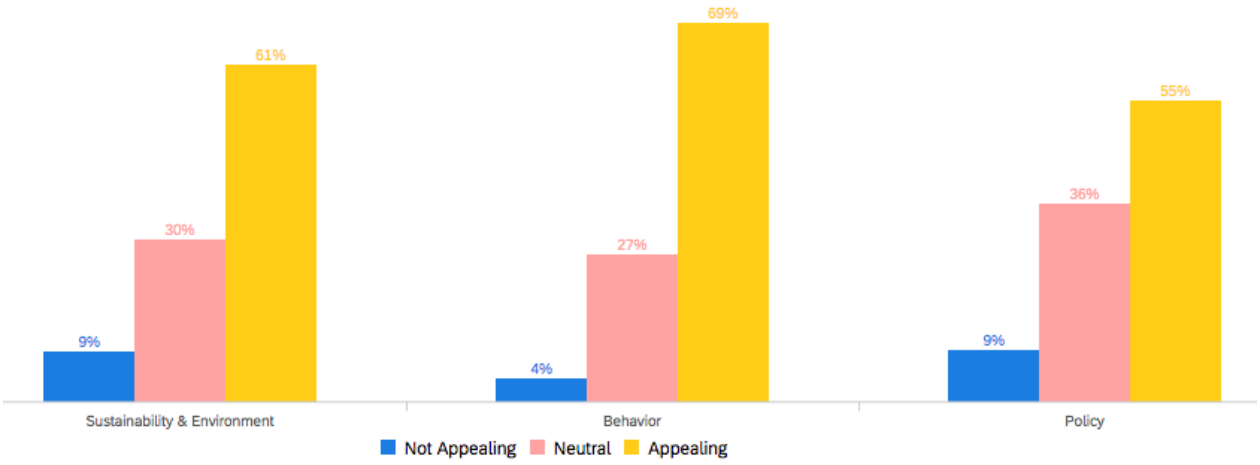
Figure 15. Interest in Proposed Courses for Master's of Data Science & Analytics

When asked in an open-ended question if there were any additional courses or topics that should be covered in a data science master's program, the following suggestions were reported:

- programming languages/software and statistical analysis platforms for big data analytics
- GIS
- Machine learning and forecasting

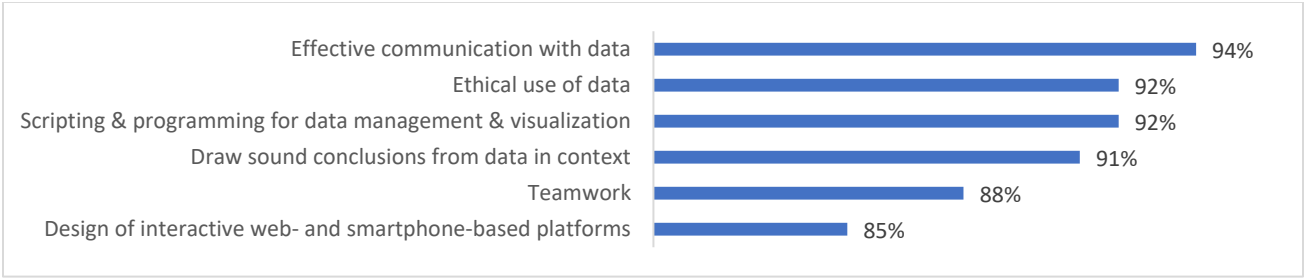
Descriptions of three proposed course themes were presented to respondents who were asked to rate if each were appealing. The "Behavior" theme was rated as most appealing (69%), followed by "Sustainability & Environment" (61%), and then "Policy" (55%). These findings are presented in Figure 16.

Figure 16. Interest in Proposed Themes for Master's of Data Science & Analytics



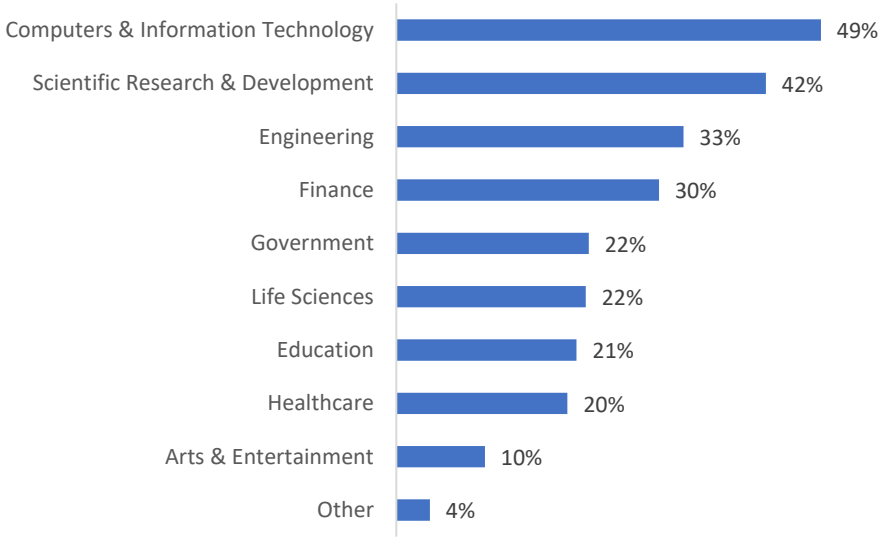
Respondents were asked to rate the value of a list of skill sets that would be learned through the proposed program. "Effective communication with data" was rated most valuable (94%), followed by "ethical use of data" (92%), and then "scripting and programing" (92%). See Figure 17 for value of proposed skills learned.

Figure 17. Value of Skills Learned through Proposed Program



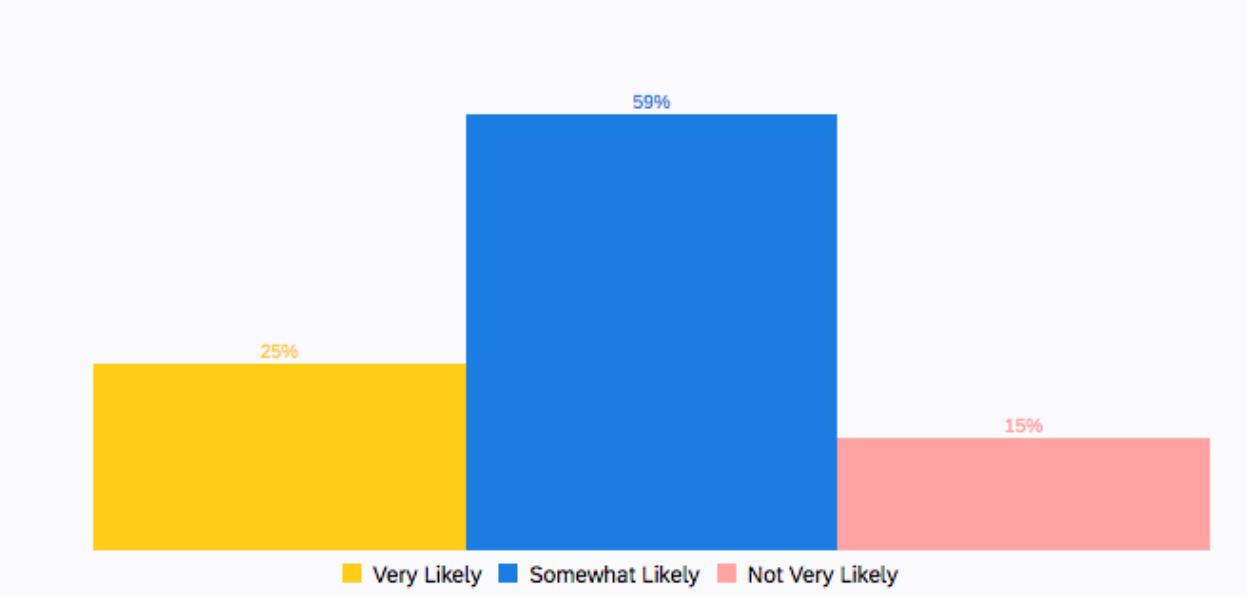
Respondents interested in the proposed Master's of Data Science & Analytics said they would hope to apply knowledge learned from the program in a diverse range of fields, as shown in Figure 18. 'Computers and information technology', and 'scientific research and development' were the most commonly cited fields.

Figure 18. Application of Proposed Degree (n=285)



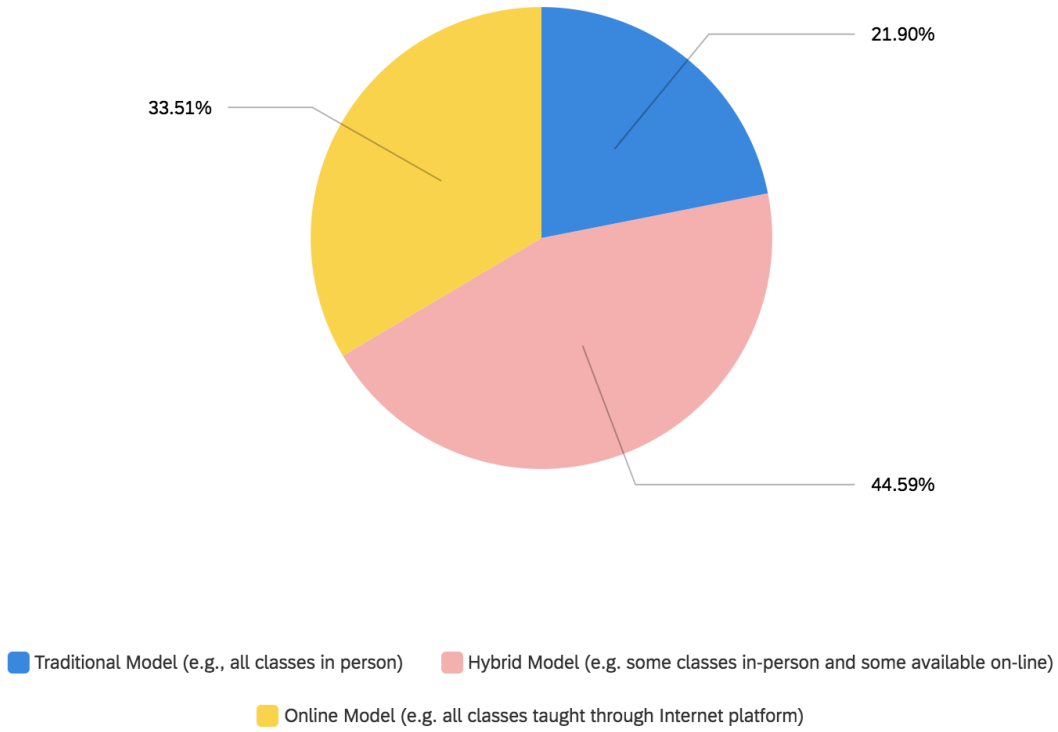
When asked their likelihood of applying to the proposed program, 59% said they were somewhat likely to apply, and 25% said they were very likely to apply, as shown in Figure 19.

Figure 19. Likelihood of Applying to Proposed Program



Although the proposed UC Merced Master's of Data Science & Analytics program is an onsite offering, it can be valuable to understand prospective students' format preferences for future planning and/or program expansion. As shown in Figure 20, 44% of respondents said they would prefer a hybrid format, 34% favor an online model, and 22% prefer a traditional on-grounds format.

Figure 20. Preferred Program Format



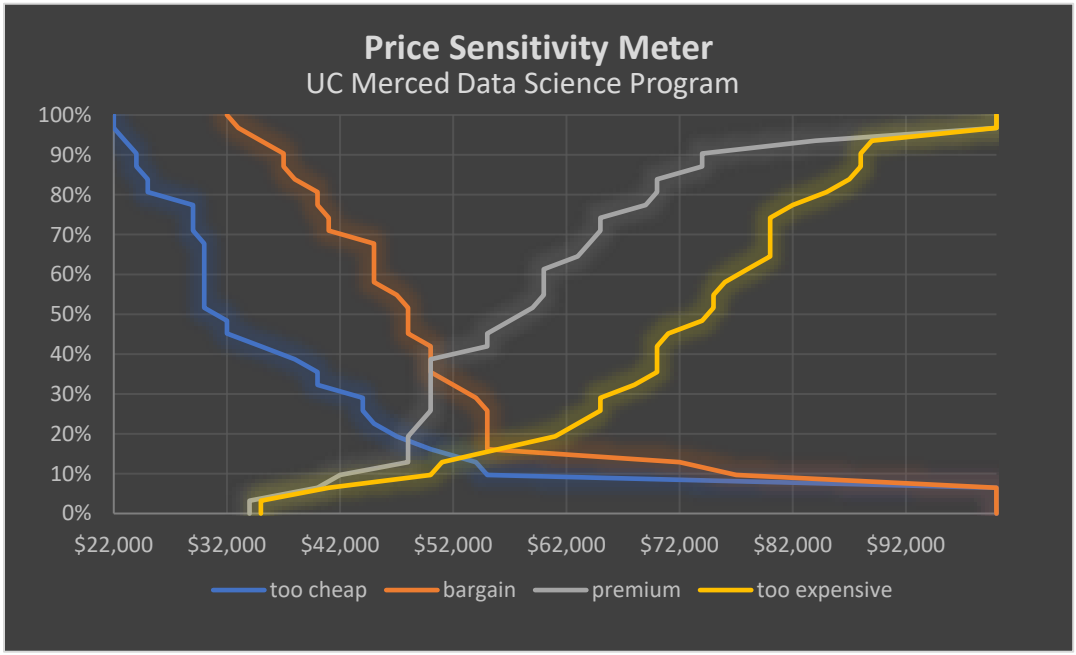
Pricing

Survey respondents were asked at what price point they would consider the proposed Master's of Data Science & Analytics program to be: 1) So inexpensive they would question the quality and not consider it, 2) A bargain—a great buy for the money, 3) Expensive but still would consider it, and 4) Too expensive to consider. They were told Master of Science in Data Science programs available at other University of California campuses ranged in price from $41,000 to $73,000. The mean statistics are as follows:

- *Too cheap to be of good quality*: **$38,323**
- *A Bargain or great buy*: **$50,645**
- *Expensive but still would consider*: **$60,097**
- *Too Expensive to consider*: **$72,677**

Responses to these four pricing questions can be graphed using the Van Westendorp model to estimate the optimal program pricing point, which is defined as the space where the four lines intersect in Figure 21.

Figure 21. Van Westendorp Pricing Sensitivity Meter



Conclusion

Over the past decade, there has been a dramatic increase in the demand for data scientists, and data science skill sets—in 2019, there were over 1 million job postings in the United States for positions that require data science skills. In response to this demand, many higher education institutions have developed, or are developing, degree programs in data science.

Prospective students who were surveyed expressed a high level of interest in the Master's of Data Science & Analytics program proposed by UC Merced. The highest rates of interest in the proposed program were expressed by the following groups: those with engineering degrees; people between the ages of 21-28; men; and current students. Prospective students said the proposed program is appealing because it would enable them to gain knowledge of data science, increase job security, and promote career advancement. Skills learned, cost, reputation are the factors most influencing prospective student decisions about enrollment. Survey respondents said if the proposed program were priced at $38k would be too cheap to be of good quality, and $72k would be too expensive to consider. Pricing sensitivity analyses suggest the optimal price point is in the $50-60k range.

APPENDIX A

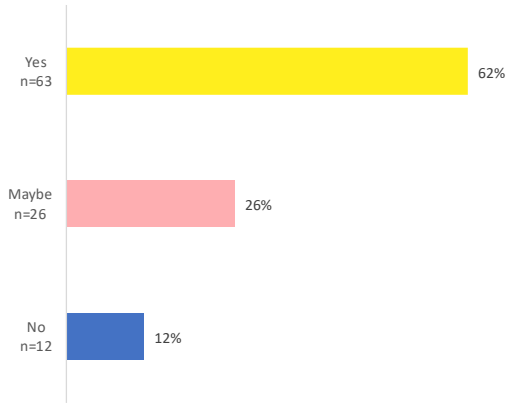
Institution	Name of Program	Location
Stanislaus State	MBA (Business Analytics Track)	Turlock, CA
Drexel University	MS Data Science	Philadelphia, PA
Illinois Tech	MS Data Science	Chicago, IL
University of Missouri	MS Data Science & Analytics	Columbia, MI
Indiana University	MS Data Science	Bloomington, IN
Johns Hopkins University	MS Data Science	Baltimore, MD
UC Berkley	Master of Information and Data Science (MIDS)	online
Syracuse University	Master's in Applied Data Science	online
Southern Methodist University	MS Data Science	online
University of Denver	MS Data Science	online
Georgia Tech	MS Analytics	online or Atlanta, GA

Background & Demographics

- As part of a market analysis for a Master's of Data Science & Analytics program proposed by UC Merced, the UC Merced Alumni Relations office distributed a survey created by the UC San Diego Center for Research and Evaluation to their alumni to measure prospective student interest and demand
- In August 2020, UC Merced Alumni Relations collected 101 survey responses from UC Merced Alumni
  - 25% of respondents were 18-24 years of age
  - 75% of respondents 25-44 years of age
  - 95% of respondents reside in California

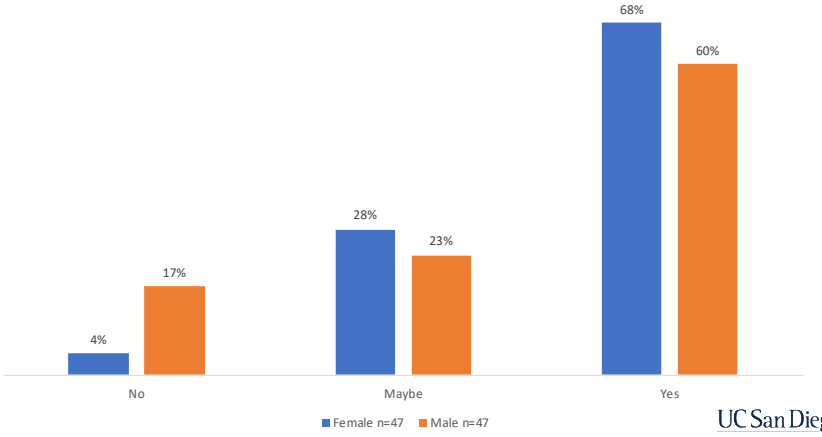
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Proposed Degree Appeal



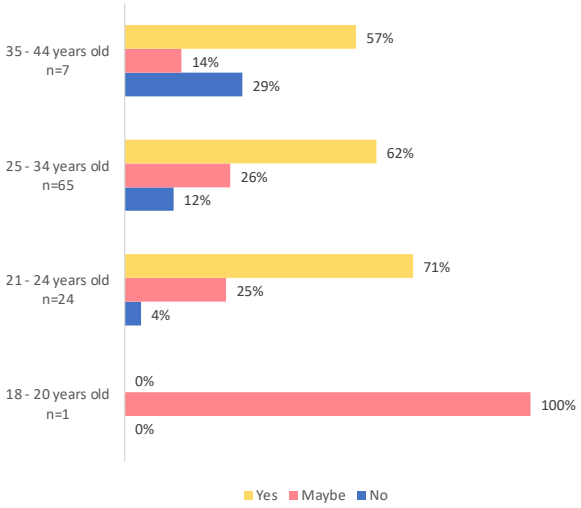
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Degree Appeal by Gender



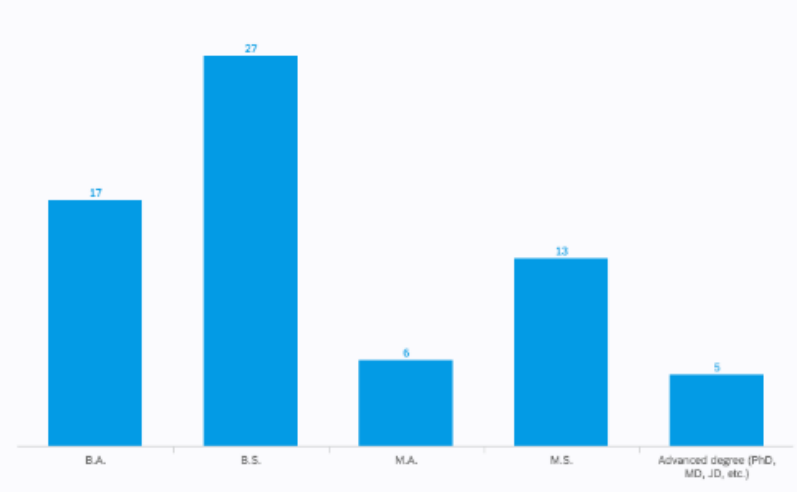
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Degree Appeal by Age



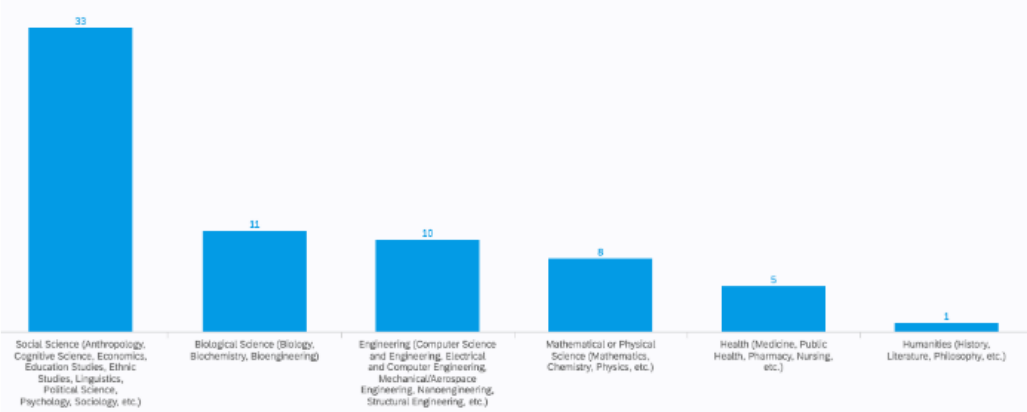
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Degree Appeal by Educational Attainment



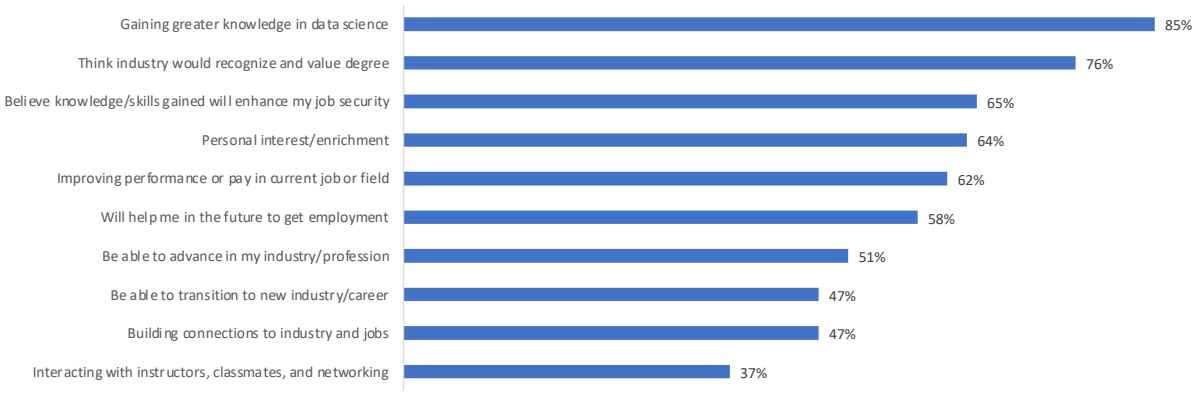
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Degree Appeal by Highest Degree Area



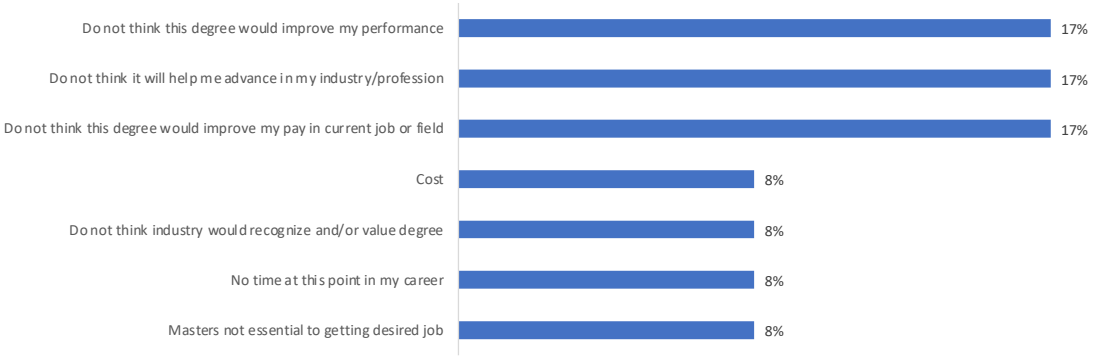
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Reasons for Proposed Program Appeal



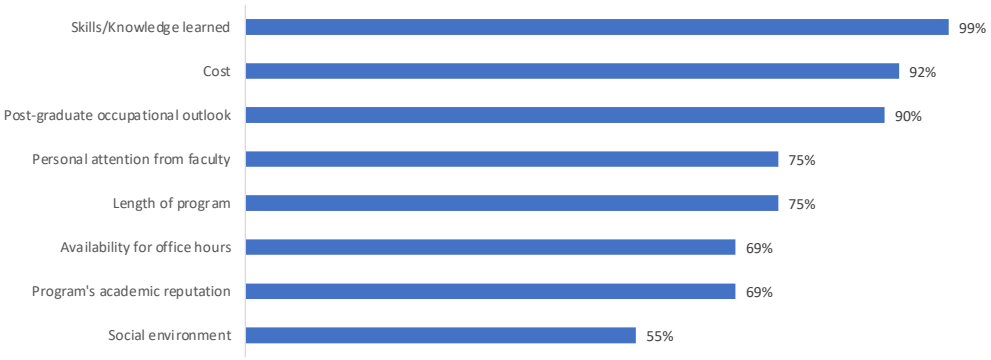
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Reasons Proposed Program is *Not* Appealing



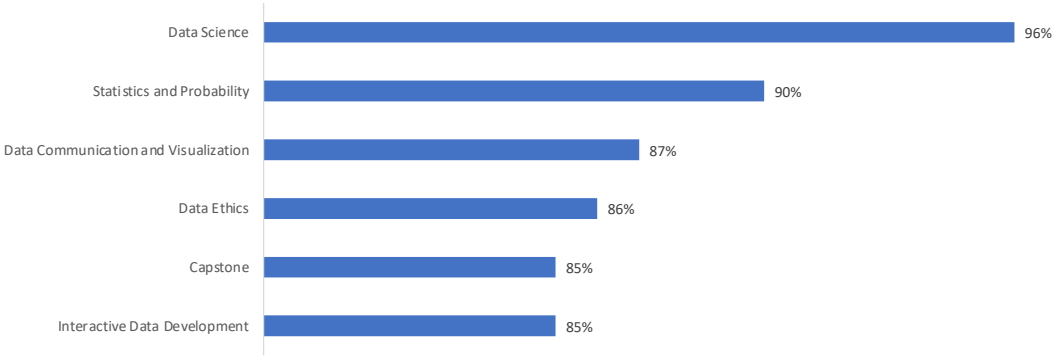
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Factors Influencing Enrollment



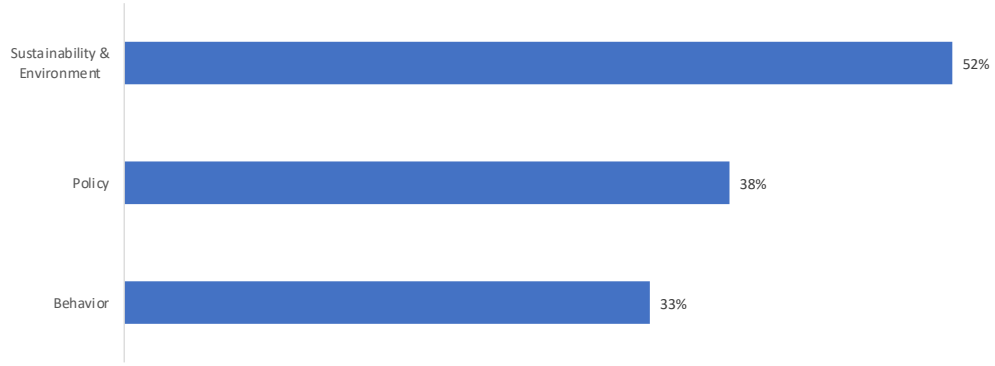
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Interest in Proposed Courses



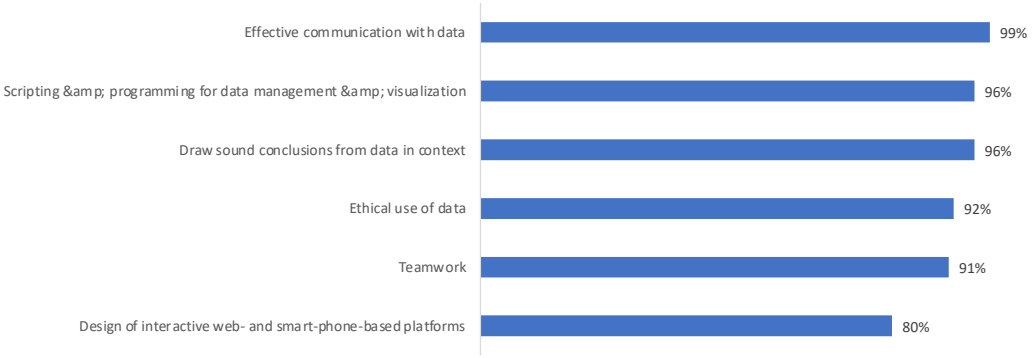
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Interest in Proposed Themes



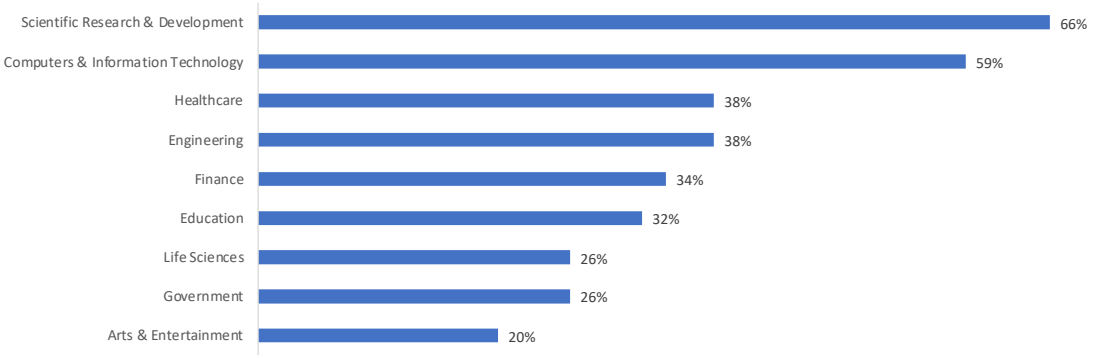
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Value of Skills Learned



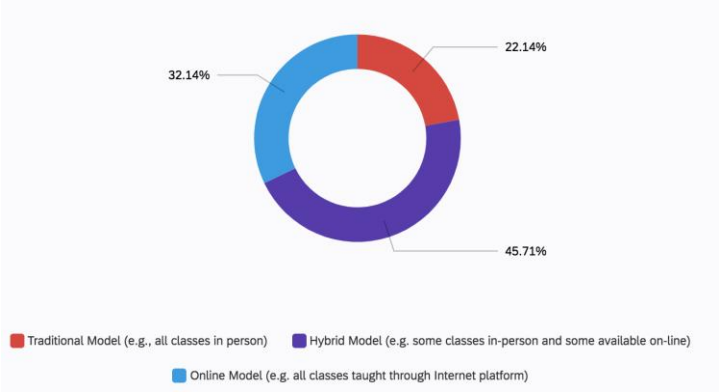
UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Application of Proposed Degree



UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

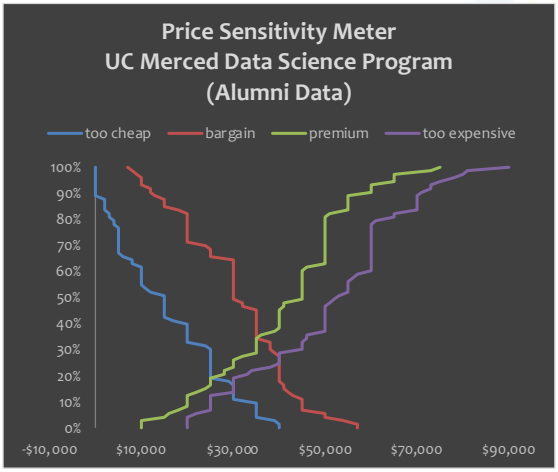
Preferred Program Format



UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Pricing Sensitivity Meter

- Mean statistics, at what price point is the proposed program:
  - *Too cheap to be of good quality*: **$14,945**
  - *A bargain or great buy*: **$30,178**
  - *Expensive but still would consider*: **$41,370**
  - *Too expensive to consider:* **$51,233**



UC San Diego EXTENSION | CR+E CENTER FOR RESEARCH + EVALUATION

Appendix F – Example Memorandum of Understanding

EXAMPLE MEMORANDUM OF UNDERSTANDING
Enrollment of Students from the Master of Data Science and Analytics

The Department and Graduate Group in [Public Health/Psychological Sciences/Environmental Systems] agrees to allow students from the Master of Data Science and Analytics to enroll in the following courses if seats are available. Students from state-supported programs shall not be denied seats in favor of self-supporting degree students.

[list courses]
PH 202: Epidemiological Methods, PH 204: Environmental Health, PH 216: Health Policy
PSY 224: Health Disparities, PSY 225: Health Risk Decision Making
ES 207: Environmental Data Analysis, ES 240: Water Resources Planning and Management

In consideration for each enrolled student, 40% of net MDSA program revenue per course will be charged. We understand that net revenue is calculated as 50% of revenue collected from student tuition (ie, revenue minus indirect cost return to the campus and funds returned to student aid), and will amount to $1,000 for a single student to enroll in a 4-unit course at the time the program starts. The fee is subject to change based on changes to net revenue for the MDSA program.

This MOU will become effective on [date] and will be in effect for 5 years, with renewal based on the mutual agreement of the graduate group chairs. Responsibility to meet the terms of this MOU shall pass to all future graduate group chairs. This MOU may be terminated before the expiration of the initial term or a renewal term upon written agreement of all parties.

Authorizing Signatures and Dates

_____	_____
Signature	Signature
_____	_____
Print Name	Print Name
_____	_____
Title and Program	Title and Program
_____	_____
Date	Date
_____	_____
Signature	Signature
_____	_____
Print Name	Print Name
_____	_____
Title and Program	Title and Program
_____	_____

Date Date

Appendix G – Internal Letters of Support

1. Haipeng Li, UC Merced Library
2. Jan Wallander, UC Merced, Graduate Group in Psychological Sciences
3. Irene Yen, UC Merced, Graduate Group in Public Health
4. Martha Conklin, UC Merced, Graduate Group in Environmental Systems
5. Brad LeVeck, UC Merced, Graduate Group in Political Science

UC MERCED LIBRARY
UNIVERSITY OF CALIFORNIA, MERCED LIBRARY • 5200 N. LAKE ROAD • MERCED, CA 95343

Date: September 28, 2020

The Library is excited to work with the proposed Gallo School and specifically to collaborate with the proposed Master of Data Science and Analytics program in planning for its establishment in the near future. On behalf of the UC Merced Library, I am pleased to write this letter in support of the Master of Data Science and Analytics proposal and agree in principle to offer Library services with the goal of developing a successful and sustainable data science program.

The Library acknowledges the great opportunity for collaboration and partnership between the Library and the Master of Data Science and Analytics. We agree on the following principles that will help us craft a specific Memorandum of Understanding as the planning moves forward:

1. Operations fee: Though minimal impact on collections is anticipated at this point, overall support from the library (e.g, SpARC's support for GIS applications during the year) is critical. Thus, the Master of Data Science and Analytics program will provide the Library an annual payment operational support (anticipated to be $5000).

2. Per student fee: The Master of Data Science and Analytics Program will pay the Library a set fee a per student for providing a Data Bootcamp, expected to be a 3-day (15 contact hour) training (anticipated to be $333 per student).

3. Scaling: It is expected that the Master of Data Science and Analytics program will start with about 10 students in the first year, and potentially grow to 40 students over 5 years. As the program grows, we share the understanding that the number of Bootcamp sessions and workshops may increase, at which point a separate plan will be developed to handle the growth. This separate agreement to scale operations may also apply to other areas of library support such as those in collections, GIS, and data literacy.

Sincerely,

Haipeng Li

University Librarian
University of California, Merced

UNIVERSITY OF CALIFORNIA

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

JAN L. WALLANDER, PH.D.,
DISTINGUISHED PROFESSOR
PSYCHOLOGICAL SCIENCES

jwallander@ucmerced.edu
http://faculty1.ucmerced.edu/jwallander/
http://psychology.ucmerced.edu

UNIVERSITY OF CALIFORNIA, MERCED
5200 NORTH LAKE RD
MERCED, CA 95343

(209) 756-5731

14 September 2020

As Chair of the Graduate Group in Psychological Sciences (PSY), I am writing to confirm the group's participation in the proposed SSGPDP Master of Data Science and Analytics (MDSA) program, which will be managed by the proposed Gallo School and is anticipated to start in Fall 2022. Specifically, PSY agrees to allow MDSA program students to enroll in the following PSY courses as long as seats are available, and in return, PSY will receive a pro-rated portion of the net revenue of the MDSA program. The specific agreement between PSY and the MDSA program will be formalized at a later date.

PSY 224: **Health Disparities**. Disease prevalence, severity, and treatment varies across sociodemographic groups. Understanding why health disparities occur is key to determining how inequalities might be alleviated. The focus of this course is on research that a) describes health disparities, b) investigates factors that explain differences, and c) proposes interventions to treat at-risk populations (Instructor: Anna Apperson)

PSY 225: **Health Risk Decision Making**. A focus on the decision-making process underlying health risk behaviors. Consideration of the role perceptions of risks/benefits, attitudes, emotions, social relationships, and the media play on health decisions, with an emphasis on decision making theories (e.g., rational choice theory, prospect theory, health beliefs model, and the theory of planned behavior). (Instructor: Anna Song).

I note that additional courses may be included by agreement in the future, and that the group does not anticipate any negative consequences to our traditional state-supported programs by including MDSA students in some of our courses.

Sincerely,

Jan L. Wallander, PhH
Chair, Graduate Group in Psychological Sciences

UNIVERSITY OF CALIFORNIA, MERCED

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

Irene H. Yen, PhD, MPH
Professor and Chair of Graduate Studies
Public Health
SSM 357B

September 21, 2020

Paul P. Maglio
School of Engineering - University of California, Merced
Professor - Ernest and Julio Gallo Management Program
Chair - Department of Management of Complex Systems
Director - Division of Management and Information

RE: Master of Data Science and Analytics program

Dear Professor Maglio:

As Chair of Graduate Studies in Public Health (PH) Department, I am writing to confirm our department's participation in the proposed self-supporting graduate professional degree program Master of Data Science and Analytics (MDSA) program, which will be managed by the proposed Gallo School and is anticipated to start in Fall 2022. Specifically, our department agrees to allow MDSA program students to enroll in the following courses as long as seats are available, and in return, our department will receive a pro-rated portion of the net revenue of the MDSA program. The specific agreement between PH and the MDSA program will be formalized at a later date.

PH 202: Epidemiological Methods. Introduces, compares, and applies conceptual frameworks, measures, study designs, and analysis approaches used in the field of epidemiology including causality, measures of disease, measures of association, study design (trials, cohort, case-control, cross-sectional and ecological), biases, screening, statistical inference, and analyzing epidemiologic data. (Instructor rotates between Professors Goldman-Mellor, Ha, and Yen).

PH 204: Environmental Health. Why the environment is essential to human health and how we analyze and act on environmental agents, factors, and conditions to improve health of people, at local, regional, and global scales. Focuses on issues pertinent to the San Joaquin Valley. (Instructor Cisneros or Bradman).

PH 216: Health Policy. Examines key health policy research topics and methods used in the field. The course is designed to provide students with an overview of health policy issues and research in the field while simultaneously developing rigorous critical analysis and research skills. Students will learn about a range of local, state, and federal policies in the U.S. to improve population health outcomes, including health care policies (e.g., Medicare, Medicaid, etc.) and public health policies to promote health. They will also learn about the health policymaking process. Research articles in leading peer-reviewed journals in the fields of health policy, public health, health services research, economics, and political science will be emphasized. (Instructor Payán).

I note that additional courses may be included by agreement in the future, and that the group does not anticipate any negative consequences to our traditional state-supported programs by including MDSA students in some of our courses.

Please feel free to contact me with any questions at iyen@ucmerced.edu.

Sincerely,

Irene H. Yen, PhD, MPH

Professor and Chair of Graduate Studies

UNIVERSITY OF CALIFORNIA

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

SCHOOL OF ENGINEERING

University of California, Merced
5200 N. Lake Road | Merced, CA 95343
PHONE: 209-228-4411
FAX: 209-228-4047

September 30, 2020

To: Paul Maglio, Chair - Department of Management of Complex Systems
From: Martha Conklin, Chair – Environmental Systems Graduate Group MHC

RE: Use of Environmental System courses in the proposed SSGPDP Master of Data Science and Analytics

As Chair of the Graduate Group in Environmental Systems (ES), I am writing to confirm the group's participation in the proposed SSGPDP Master of Data Science and Analytics (MDSA) program, which will be managed by the proposed Gallo School and is anticipated to start in Fall 2022. Specifically, ES agrees to allow MDSA program students to enroll in the following ES courses if seats are available, and in return, ES will receive a pro-rated portion of the net revenue of the MDSA program. The specific agreement between ES and the MDSA program will be formalized at a later date.

ES 207: Environmental Data Analysis. The objective of this class is to provide students with probabilistic and statistical methods to analyze environmental data. This class emphasizes both theoretical and applied aspects of data analysis methods. Weekly lab exercises are from environmental applications. Topics include distribution, hypothesis test, linear regression, multiple regression, uncertainty analysis, outlier detection, sample design, and spatial and temporal data analysis. (Instructor TBA).

ES/MIST 232 Applied Climatology. Spatial and temporal patterns in climate and their association with land surface characteristics and processes. Methods for exploiting these for hypothesis testing, modeling, and forecasting. Applications include seasonal forecasting, ecological modeling, and analysis of processes such as flooding and wildfire. (Instructor TBA).

ES 240: Water Resources Planning and Management. Basic concepts of and issues in water resources management, water resources planning, institutional and policy processes. Quantitative analytical methods in water resources planning and management; introduction to systems analysis, multi-objective planning, and risk assessment. Design project. Graduate requirements include preparation of a detailed case analysis. (Instructor TBA).

ES 260: Sustainable Energy. Current systems for energy supply and use. Renewable energy resources, transport, storage, and transformation technologies. Technological opportunities for

improving end-use energy efficiency. Recovery, sequestration, and disposal of greenhouse gases from fossil-fuel combustion. Graduate requirements include preparation of a detailed case analysis. (Instructor TBA).

ES 292: Life Cycle Assessment. Life cycle assessment (LCA) is a tool used across fields to determine the cradle-to-grave environmental impacts of products and systems. The course will cover how to mathematically define the life cycles of products and systems, perform an LCA, and interpret LCA results and evaluate them within the context of the scientific literature. Students in the course will individually conduct a complete life cycle assessment with a literature review, sensitivity analysis, and uncertainty analysis using available data and impact assessment methods with guidance from the instructor.. (Instructor TBA).

I note that additional courses may be included by agreement in the future, or equivalent courses may be substituted due to faculty availability. The group does not anticipate any negative consequences to our traditional state-supported programs by including MDSA students in some of our courses.

UNIVERSITY OF CALIFORNIA

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

Brad LeVeck
Associate Professor
Department of Political Science
(510) 604-0893
bleveck@ucmerced.edu

UNIVERSITY OF CALIFORNIA, MERCED
5200 NORTH LAKE ROAD
MERCED, CA 95343

26 August 2022

As Chair of the Graduate Group in Political Science (POLS), I am writing to confirm the group's participation in the proposed SSGPDP Master of Data Science and Analytics (MDSA) program, which will be managed by the proposed Gallo School and is anticipated to start in Fall 2022. Specifically, POLS agrees to allow MDSA program students to enroll in the following POLS courses as long as seats are available, and in return, POLS will receive a pro-rated portion of the net revenue of the MDSA program. The specific agreement between POLS and the MDSA program will be formalized at a later date.

POLI 213: Experimental Methods in Political Science. This course is intended to provide students with an understanding of experimental methods in Political Science. The first portion of the semester will emphasize concepts and tools from the experimentalist's toolbox with a strong focus on causal inference, external and internal validity, and choosing subjects and subject's motivations. Later weeks will focus on issues and challenges to specific types of Political Science experiments including survey experiments, laboratory experiments, physiological experiments, and field experiments. (Instructor: Elaine Denny)

POLI 219: Behavioral Game Theory - Special Topics in Political Science Methodology. Behavioral game theory uses a variety of methods to understand the connection between game theory and human behavior. Game theory is the predominate method for mathematically modeling and analyzing strategic interactions in political science and economics. However, many conventional modeling assumptions are based on introspection and guesses, rather than careful observation of how people actually play in games" (Camerer 2003, pg. 3). Behavioral game theory tries to figure out when these simplifying assumptions provide insight into human behavior. It also tries to develop models that can explain behavioral deviations from the traditional assumptions and tries to test the additional hypotheses that these models generate. Throughout this course, students will learn how experiments and other empirical methods have been used to answer important questions related to game theory and game theoretic models. Doing so will hopefully give students a better understanding of how they might use similar experimental tools in their own research. (Instructor: Brad LeVeck)

POLI 219: Model Based Inference - Special Topics in Political Science Methodology. This is an advanced graduate methodology class which will focus on the use of empirical models for statistical description and prediction with social science applications. The class takes a "learning" approach to modeling, covering GLMs and machine learning. this class follows from the probability, regression, and causal inference classes in the political science methodology sequence. (Instructor: Tesalia Rizzo)

POLI 251: Political Cognition. The seminar provides students with an overview of political psychology. How do we think and feel about politics? What preferences do we have and what motivates us to become engaged? How does the political context shape our political psychology? And how can we answer these questions? We will draw on research from political science, including the study of American politics, comparative politics, and international relations, as well as cognate disciplines like economics, psychology, sociology to answer these and other questions. (Instructor: Christopher Ojeda)

POLI 224: Subnational Politics. This seminar provides an overview of major debates in research on subnational politics. We examine origin and evolution of subnational political environments. We trace the historical development of local government institutions, analyze coalitions, investigate distributions of power, investigate the process and consequences of suburbanization, and examine local development and education policy. The primary goals of the course are to familiarize students with the principal questions asked by scholars in this subfield, the methodological approaches employed, and the avenues available for future research. Methodological diversity will be emphasized. (Instructor: Jessica Trounstine)

I note that additional courses may be included by agreement in the future, and that the group does not anticipate any negative consequences to our traditional state-supported programs by including MDSA students in some of our courses.

Sincerely,

Brad LeVeck
Associate Professor and Director of Graduate Studies
Department of Political Science,
University of California, Merced
bleveck@ucmerced.edu
(510) 604-0893

Appendix H – External Community Letters of Support

1. Jim Spohrer, IBM Research
2. Greg Gearheart, State Water Resources Control Board
3. Nicholas Martorano, California Water Quality Monitoring Council
4. Geri Miller, Esri
5. Wayne O. Miller, Lawrence Livermore National Laboratory
6. David Bell, Research Institute for Advanced Computer Science
7. Cicely Muldoon, United States Department of the Interior, National Park Service

August 29, 2020

Re: Letter of Support for Proposed Master of Data Science and Analytics at UC Merced

Dear Dr. Maglio,

Thank-you for presenting virtually at the AHFE HSSE conference last month on AI Ethics and Service. Thank-you also for following up with the description of the proposed UC Merced Master of Data Science and Analytics document and making time to discuss more of the details of the degree program. Previously, my IBM colleague Daryl Pereira spoke highly of related work, when he presented at a UC Merced Data Science Summit – so it is great to continue growing the relationship between IBM and Merced in this important area of Data Science.

The one-year intensive master degree program is well designed to meet the growing need of industry to transform massive datasets into business insights/advantage in an ethical manner. Furthermore, because IBM hires tens of thousands of new employees each year (both as interns and full-time hires) and value diversity and inclusion as part of our core principles, the diversity in the UC Merced student populations is also a good match for IBM. In short, IBM would very much welcome the resumes of students who successful complete the intensive one-year program as part of the applicant pool for both entry-level Data Science and Junior Data Scientist Apprentice positions at IBM.

In sum, I whole-heartedly support and encourage proposal. On the second page of this support letter, I have included links to IBM's policies and procedures to ensure ethical use of technology. We work diligently to review, revise, and improve these processes – and as noted in my first sentence above grateful for conference presentation about your research on this important topic.

Sincerely yours,

James C. Spohrer

Dr. James ("Jim") C. Spohrer
Director, Open Cognitive Technologies, Cognitive Applications/Digital Ecosystem Group
IBM Research – Almaden, 650 Harry Road San Jose, CA 95120
spohrer@us.ibm.com, 408-927-1928 (office)
spohrer@gmail.com, 408-829-3112 (mobile)
Innovation Champion: http://service-science.info/archives/2233

IBM's Policies and Principles
https://www.ibm.org/responsibility/policies

IBM's Human Rights Principles
https://www.ibm.org/responsibility/policies#humanRights

IBM's Cognitive Principles
https://www.ibm.org/responsibility/policies#cognitivePrinciples

IBM's Principles for Trust and Transparency
https://www.ibm.com/blogs/policy/wp-content/uploads/2018/05/IBM_Principles_OnePage.pdf

IBM's Principles for Ethical Design of AI Systems
https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf

IBM's AI Design Ethics Overview
https://www.ibm.com/design/ai/ethics/

IBM Values
https://www.ibm.com/ibm/values/us/

IBM's Business Conduct Guidelines (BCG)
https://www.ibm.com/investor/att/pdf/BCG_accessible_2019.pdf

Partnership for AI - Tenets
https://www.partnershiponai.org/tenets/

Linux Foundation AI - Trusted AI Committee
https://wiki.lfai.foundation/display/DL/Trusted+AI+Committee
https://developer.ibm.com/blogs/ibm-linux-foundation-ai-trusted-workflows/

IBM Research - Science for Good
https://www.research.ibm.com/science-for-social-good/

IBM Research Partnership for AI and Sustainable Development
https://www.research.ibm.com/science-for-social-good/#overview
https://www.undp.org/content/undp/en/home/blog/2018/ai-and-the-future-of-our-work.html

IBM Foundation Partnership to Fight Human Trafficking
https://www.ibm.org/initiatives/human-trafficking/ht-1

IBM Contact Points - Worldwide
https://www.ibm.com/planetwide/

IBM Contacts - Business Functions
https://newsroom.ibm.com/contacts

IBM Contact - Government and Regulatory Affairs
https://www-03.ibm.com/press/us/en/contact/46097.wss?topic=491

IBM AI Ethics Board Co-Chairs
https://newsroom.ibm.com/Christina-Montgomery
https://www.ibm.com/blogs/policy/francesca-rossi-ai/

CALIFORNIA Water Boards

GAVIN NEWSOM
GOVERNOR

JARED BLUMENFELD
SECRETARY FOR
ENVIRONMENTAL PROTECTION

State Water Resources Control Board

September 3, 2020

Prof. Dan Hicks
Assistant Professor
Cognitive and Information Sciences
University of California, Merced

Dear Dan,

It is with great enthusiasm that I write this letter of support for the proposed Master of Data Science and Analytics at UC Merced. I believe the proposed will be of great benefit to our communities, particularly those in Central California, by training professionals with skills related to data science and analytics.

I am the Director of the Office of Information Management and Analysis at the CA Water Boards. We employ about 2300 staff around the state in our mission to protect water for all Californians. Our current and future work requires better data science skills for all job classifications, most notable those whom enter our workforce as scientists, engineers, geologists and analysts. In addition, my office employs data scientists in a new series of classifications called the Research Data series. Having graduates of our state universities come to work with ready to deploy data science skills will make a huge difference in the level of services we can provide. We also hire interns and fellows and do civic data engagement with emerging academics, so we are very excited to hear about this new program.

I fully support the establishment of the professional Master of Data Science and Analytics program at the University of California at Merced.

Sincerely,

Greg Gearheart, Deputy Director
Office of Information Management and Analysis

E. JOAQUIN ESQUIVEL, CHAIR | EILEEN SOBECK, EXECUTIVE DIRECTOR

1001 I Street, Sacramento, CA 95814 | Mailing Address: P.O. Box 100, Sacramento, CA 95812-0100 | www.waterboards.ca.gov

California Water Quality Monitoring Council

Jared Blumenfeld
Secretary for Environmental Protection

Nicholas Martorano, Director
Mailing Address: Nick Martorano, Monitoring Council Director
1001 I Street, 19th Floor, 19-65A • P.O. Box 100
Sacramento, California • 95812-0100
(916) 341-5514
mywaterquality.ca.gov

Gavin Newsom
Governor

Monitoring Council Members
Karen Mogus
Environmental Protection Agency
Chad Dibble
Natural Resources Agency
Andrew Altevogt
Division of Drinking Water
Grant Sharp
Stormwater
Shelly Walther
Publicly Owned Treatment Works
Amrith Gunasekara
Agriculture
Helen Fitanides
Citizen Monitoring
Ray Hiemstra
Public
Stephen Weisberg
Scientific Community
Peter Vroom
Water Supply

Monitoring Council Staff
Nicholas Martorano
Director

September 8, 2020

Prof. Dan Hicks
Assistant Professor
Cognitive and Information Sciences
University of California, Merced

It is with great enthusiasm that I write this letter of support on behalf of the California Water Quality Monitoring Council (Monitoring Council) for the proposed Master of Data Science and Analytics at UC Merced. The Monitoring Council believes the proposed degree program will be of great benefit to our communities, particularly those in Central California, by training professionals with skills related to data science and analytics.

The proposed program's focus on data science from a human perspective seems particularly appropriate to support urgent needs of industry and government to understand and manage pressing local and global problems. The Monitoring Council is tasked by its founding legislation to help coordinate and integrate water quality and ecosystem health monitoring across California. With that the integration, management and communication of data and results are paramount. The Monitoring Council sees a strong potential for graduates to be in high demand across sectors that are represented on the Monitoring Council and will provide tremendous value to all of California.

The Monitoring Council fully supports the establishment of the professional Master of Data Science and Analytics program at the University of California at Merced.

Sincerely,

Nicholas Martorano
Director, California Water Quality Monitoring Council

Esri Inc.
380 New York Street
Redlands, CA 92373
September 16, 2020

Dear Dr. Maglio:

I am writing to express our support for the establishment of the newly proposed masters program in *Data Science and Analytics* at the University of California - Merced, E&J Gallo School of Management.

Many of the challenges society faces in this century, will be influenced by increasingly complex global issues. Experience shows we will require an understanding of international perspectives, earth systems, emerging technologies, and the geographic relationships between humans and their planet, as we strive to develop science-based policy and business solutions.

This understanding is critical if corporations, public agencies and not for profit organizations, are to flourish with a progressive emphasis on sustainable means and the employment of methods, focused on customer care, workforce wellness, and fair, efficient impact across supply chains.

Success will require understanding how to acquire, organize, and analyze staggering amounts of data – and how to efficiently retrieve the data we need from the massive resource available from public, government and corporate domains. At Esri, we pioneer methods for just this task and strive to make data comprehensible through innovative analytics and graphic representation.

I believe your proposed degree program will be relevant and of significant benefit to our broad and diverse California community, by training professionals with the ever-evolving skills required for this important work.

Your program's emphasis on data science from the human perspective is timely. It recognizes the relationship of data to useful service. True from both local and global viewpoints. At Esri, through interaction with our tens of thousands of clients around the world, we see firsthand, the daily interplay between corporate goals, government policy, public health, environmental stewardship, science and education. And much more. Naturally, we understand the need to develop the necessary workforce for this industry through cutting edge academic preparation. This is what you propose.

For all of these reasons, we at Esri support the establishment of the professional Master of Data Science and Analytics program at the University of California, Merced and look forward to observing the positive impact of your graduates.

Sincerely,

Geri Miller

Geri Miller
Education Sector Director
Global Business development
Esri

ESRI • 380 New York St., Redlands, CA 92373-8100, USA • TEL 909-793-2853 • FAX 909-793-5953 • E-MAIL info@esri.com • WEB www.esri.com

HIGH PERFORMANCE COMPUTING
INNOVATION CENTER

7000 East Avenue, L-780, Livermore, CA 94550
hpcinnovationcenter.llnl.gov
925.424.8000

September 23, 2020

Professor Paul P. Maglio
Chair, Department of Management of Complex Systems
UC Merced

Dear Prof. Maglio,

I'm writing in full support of your proposed Master of Data Science and Analytics (MDSA) degree program at UC Merced. It will fill a strong need for the U. S. workforce in general and particularly for my own institution, the Lawrence Livermore National Laboratory (LLNL).

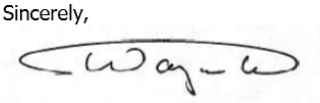
Data Science and Analytics is a technical path that has had explosive growth over the last decade due to the strong emergence and acceptance of data analytics, machine learning, and AI by the private sector. As you are aware, the tools and techniques for applying these methods have matured to the point of significant practical utility…they have successfully moved from research to practice. The broad utility of these methods means that they have been adopted in all sectors, from finance to medical research. This broad adoption is the driving need for a trained workforce in this technical path, and your Master's program is a timely answer to this need.

In particular at LLNL, we have mirrored this growth and adoption of data analytical methods in both practical and R&D applications. We apply data analytics for robust decision making, and also develop new approaches to data science that will move the field forward. This has public evidence in our Data Science Institute (DSI)[1], which serves as an outreach organization for public engagement between LLNL and other institutions such as UC Merced. In fact, the DSI has a particularly strong engagement with the UC System as a whole, including an annual workshop for UC students and faculty, and the Data Science Summer Institute (DSSI) for students that has been highly popular.

LLNL looks to UC Merced as a long-term partner for both academic engagement and as a pipeline for new employees from your student body. We have many 1:1 faculty engagements on cooperative R&D activities and support students through internships and fellowships for undergraduate and graduate studies. Your MDSA degree program will strengthen this bond by creating a new formal program in a target area for LLNL, and I expect that many students and faculty who are part of the MDSA program will find opportunities to work with LLNL staff scientists in data analytics, to our mutual benefit.

In summary I strongly support the creation of your MDSA degree program as it answers a need for workforce development in this technical area and will only strengthen the bond between UC Merced and LLNL. I will look forward to working with you soon when the program is opened.

Sincerely,



Wayne O. Miller
Deputy Director, HPC Innovation Center
Lawrence Livermore National Laboratory
Miller99@llnl.gov

[1] https://data-science.llnl.gov/

Lawrence Livermore National Laboratory

RESEARCH INSTITUTE FOR ADVANCED COMPUTER SCIENCE
UNIVERSITIES SPACE RESEARCH ASSOCIATION
615 National Ave, Suite 220
Mountain View CA 94043
Phone: 650-575-4362 * Fax: 650-966-5021

Date: September 25, 2020

To Whom It May Concern:

I am writing to confirm my interest in collaborating with Prof. Lace Padilla of the University of California, Merced, in support of a new UC Merced Master of Data Science and Analytics Program. In particular, my organization is interested in collaborating with her Global Good Studio class, working to engage students involved in the course through internships and course projects that align with research projects we are conducting with our research sponsors at the NASA Ames Research Center and the U.S. Geological Survey's National Innovation Center, the USGS Earthquake Science Center, and the USGS Geology, Minerals, Energy, and Geophysics Science Center.

This will build well upon a collaboration that started earlier this year with Prof. Padilla when a student team in her course implemented one of our research projects on "Predictive Modeling of Flood Susceptibility" http://space.ucmerced.edu/USRA. This will also build well upon our current scope of work under the NASA Academic Mission Services (NAMS) contract we manage at NASA's Ames Research Center, which includes scope for collaborative curriculum development with faculty at universities in areas that building upon the workforce development needs of NASA.

My organization is actively working to collaboratively develop new data science curriculum with faculty at multiple universities, supporting development and sharing of open curriculum within the academic community. As one example, my organization collaboratively developed a new course on "Data Science in Aviation" which is being taught in the Department of Civil & Environmental Engineering at the University of California, Berkeley this fall. This new Master's program at UC Merced will provide new opportunities for collaborative curriculum development and student engagement, and we are very committed to collaborating with Prof. Padilla in these areas.

Best Regards,

Dr. David Bell, USRA Director and Chief Technologist
USRA Research Institute for Advanced Computer Science (RIACS)
NASA Academic Mission Services (NAMS) Contract

United States Department of the Interior

NATIONAL PARK SERVICE
Yosemite National Park
P. O. Box 577
Yosemite, California 95389

IN REPLY REFER TO:
7.B (YOSE-SUPT)

SEP 30 2020

Dr. Paul Maglio
Professor, E&J Gallo Management Program
School of Engineering
University of California
5200 North Lake Road
Merced, CA 95343

Dear Dr. Maglio:

I am writing to recommend approval of University of California Merced's current initiative to establish a master's degree in the field of data science and analytics, within the Gallo Management Program.

Over the last three decades, Yosemite National Park has faced a variety of extraordinary management challenges. Park visitation has risen about 30% since the early 1990's and in 2016, topped five million visitors. The majority of visitor impact is concentrated in Yosemite Valley, where it overwhelms the traffic infrastructure, creating untenable delays, erosion of visitor experience, and unacceptable safety challenges for emergency operations. Changes in climate and drought have precipitated massive forest die-off and increased the severity of fire seasons to record and catastrophic scales. The recent incidence of COVID-19 has forced on-the-fly solutions to visitor access, while necessitating novel solutions for the mitigation of risk. All these challenges and many more call out for science-based policy solutions.

I believe the proposed degree program will be of great benefit to our work in Yosemite, and to that of our allied partners in resource management in the surrounding national parks and forests of Central California, by training professionals with much needed skills related to data science and analytics.

Yosemite is designated as a UNESCO World Heritage Site. It is truly an international destination. The proposed program's focus on data science from a human perspective seems particularly appropriate to support the urgent needs of our work in the public sector in order to understand and manage both our pressing local issues and their potentially global ramifications. Prior to my current assignment in Yosemite, I served in similar roles on California's northern coast and in the Bay Area; from these collective perspectives, I see strong potential for graduates to be in demand beyond our region.

On behalf of my colleagues in California's family of national parks, I fully encourage the establishment of the professional Master of Data Science and Analytics program at the University of California at Merced.

Sincerely,

Cicely Muldoon
Acting Superintendent

Appendix I – Requests for Letters from Similar UC Programs

Revised Proposal Request: July 2022

1. Claire Tomlin, UC Berkeley, EECS
2. Steve Gaines, UC Santa Barbara, Bren School
3. Bill Lin, UC San Diego, ECE
4. Alin Deutsch, UC San Diego, CSE

Original Request: September 2020

1. Vladimir Stojanovic, UC Berkeley, EECS
2. Steve Gaines, UC Santa Barbara, Bren School
3. Bill Lin, UC San Diego, ECE
4. Alin Deutsch, UC San Diego, CSE

No letters were received.

From: Paul Maglio
Sent: Wednesday, July 27, 2022 2:31 PM
To: tomlin@eecs.berkeley.edu
Cc: Alvin Cha; Yolanda Kyle
Subject: Re: Request for Comments: UC Merced Proposed Professional Master of Data Science and Analytics
Attachments: Data Science Self-Supporting Program Proposal - Clean Draft 20220726.pdf

Professor Tomlin,

Nearly two years ago, we reached out for input on our proposal for a new SSPGDP, Master of Data Science and Analytics - to the prior chair of your Master of Engineering - Data Science/Systems Concentration program. We did not hear back.

Our proposal underwent campus review, and we are finally ready to resubmit a revised version at the start of fall semester. But first, we are circulating our revision again to chairs of related programs across the system - per my original note below - to ask for your feedback, as required by CCGA.

If you wish to provide feedback, it would be most helpful if we received it by August 17.

Please let me know if you have any questions — I would be happy to chat if necessary.

Thanks.

-Paul

———————————————

Paul P. Maglio - School of Engineering - University of California, Merced
Professor - Ernest and Julio Gallo Management Program
Chair - Master of Management Professional Degree Program
Director - Division of Management and Information

Admin - Yolanda Kyle - ykyle@ucmerced.edu - 209-228-3479

> On Sep 14, 2020, at 3:52 PM, Paul Maglio <pmaglio@ucmerced.edu> wrote:
>
> Dear Professor Stojanovic,
>
> At UC Merced we are in the process of proposing a new graduate program leading to Master of Data Science and Analytics. In accordance with the review policy established by the systemwide Coordinating Committee of Graduate Affairs (CCGA), I am providing you, as the Chair of an existing comparable program (Master of Engineering - Data Science/Systems Concentration), with a copy of the current draft of our proposal. We would be very grateful for any feedback you may wish to offer us, so that the proposal may be made as strong as possible before submission.
>
> As background, please understand that the format and contents of the proposal follow the required outline found in the CCGA Handbook, and that internal and external reviewers will later be asked to address the following four points when examining our final submission:

– Quality and academic rigor of the program
– Adequacy of the size and expertise of faculty to administer the program
– Adequacy of the facilities and budgets
– Applicant pool and placement prospects for the graduates

If you wish to provide feedback, it would be most helpful if we received it within two weeks, as we expect to submit the proposal for campus review in early October.

Please let me know if you have any questions — I would be happy to chat if necessary.

Thanks.

-Paul

_____

Paul P. Maglio - School of Engineering - University of California, Merced
Professor - Ernest and Julio Gallo Management Program
Chair - Department of Management of Complex Systems
Director - Division of Management and Information

Admin - Yolanda Kyle - ykyle@ucmerced.edu - (209) 228-3479

<DRAFT_Data Science Self-Supporting Program Proposal 20200914.pdf>

From: Paul Maglio
Sent: Wednesday, July 27, 2022 2:31 PM
To: gaines@bren.ucsb.edu
Cc: Alvin Cha; Yolanda Kyle
Subject: Request for Comments: UC Merced Proposed Professional Master of Data Science and Analytics
Attachments: Data Science Self-Supporting Program Proposal - Clean Draft 20220726.pdf

Dear Dean Gaines,

Nearly two years ago, we reached out for your input on our proposal for a new SSPGDP, Master of Data Science and Analytics. Our proposal underwent campus review, and we are finally ready to resubmit a revised version at the start of fall semester. But first, we are circulating our revision again to chairs of related programs across the system - per my original note below - to ask for your feedback, as required by CCGA.

If you wish to provide feedback, it would be most helpful if we received it by August 17.

Please let me know if you have any questions — I would be happy to chat if necessary.

Thanks.

-Paul

———————————————————

Paul P. Maglio - School of Engineering - University of California, Merced
Professor - Ernest and Julio Gallo Management Program
Chair - Master of Management Professional Degree Program
Director - Division of Management and Information

Admin - Yolanda Kyle - ykyle@ucmerced.edu - 209-228-3479

Begin forwarded message:

From: Paul Maglio <pmaglio@ucmerced.edu>
Subject: Request for Comments: UC Merced Proposed Professional Master of Data Science and Analytics
Date: September 14, 2020 at 3:52:40 PM PDT
To: "gaines@bren.ucsb.edu" <gaines@bren.ucsb.edu>
Cc: Rado Lee <rlee87@ucmerced.edu>, Yolanda Kyle <ykyle@ucmerced.edu>

Dear Dean Gaines,

At UC Merced we are in the process of proposing a new graduate program leading to Master of Data Science and Analytics. In accordance with the review policy established by the systemwide Coordinating Committee of Graduate Affairs (CCGA), I am providing you, as the Chair of an existing comparable program (Master of Environmental Data Science), with a copy of the current draft of our proposal. We

would be very grateful for any feedback you may wish to offer us, so that the proposal may be made as strong as possible before submission.

As background, please understand that the format and contents of the proposal follow the required outline found in the CCGA Handbook, and that internal and external reviewers will later be asked to address the following four points when examining our final submission:

– Quality and academic rigor of the program
– Adequacy of the size and expertise of faculty to administer the program
– Adequacy of the facilities and budgets
– Applicant pool and placement prospects for the graduates

If you wish to provide feedback, it would be most helpful if we received it within two weeks, as we expect to submit the proposal for campus review in early October.

Please let me know if you have any questions — I would be happy to chat if necessary.

Thanks.

-Paul

———————————————

Paul P. Maglio - School of Engineering - University of California, Merced
Professor - Ernest and Julio Gallo Management Program
Chair - Department of Management of Complex Systems
Director - Division of Management and Information

Admin - Yolanda Kyle - ykyle@ucmerced.edu - (209) 228-3479

From: Paul Maglio
Sent: Wednesday, July 27, 2022 3:10 PM
To: billlin@ucsd.edu
Cc: Alvin Cha; Yolanda Kyle
Subject: Request for Comments: UC Merced Proposed Professional Master of Data Science and Analytics
Attachments: Data Science Self-Supporting Program Proposal - Clean Draft 20220726.pdf

Dear Professor Lin,

Nearly two years ago, we reached out for your input on our proposal for a new SSPGDP, Master of Data Science and Analytics. Our proposal underwent campus review, and we are finally ready to resubmit a revised version at the start of fall semester. But first, we are circulating our revision again to chairs of related programs across the system - per my original note below - to ask for your feedback, as required by CCGA.

If you wish to provide feedback, it would be most helpful if we received it by August 17.

Please let me know if you have any questions — I would be happy to chat if necessary.

Thanks.

-Paul

Paul P. Maglio - School of Engineering - University of California, Merced
Professor - Ernest and Julio Gallo Management Program
Chair - Master of Management Professional Degree Program
Director - Division of Management and Information

Admin - Yolanda Kyle - ykyle@ucmerced.edu - 209-228-3479

> On Sep 14, 2020, at 3:52 PM, Paul Maglio <pmaglio@ucmerced.edu> wrote:
>
> Dear Professor Lin,
>
> At UC Merced we are in the process of proposing a new graduate program leading to Master of Data Science and Analytics. In accordance with the review policy established by the systemwide Coordinating Committee of Graduate Affairs (CCGA), I am providing you, as the Chair of an existing comparable program (Master of Science in ECE - Data Science/Mcachine Learning Focus), with a copy of the current draft of our proposal. We would be very grateful for any feedback you may wish to offer us, so that the proposal may be made as strong as possible before submission.
>
> As background, please understand that the format and contents of the proposal follow the required outline found in the CCGA Handbook, and that internal and external reviewers will later be asked to address the following four points when examining our final submission:

– Quality and academic rigor of the program
– Adequacy of the size and expertise of faculty to administer the program
– Adequacy of the facilities and budgets
– Applicant pool and placement prospects for the graduates

If you wish to provide feedback, it would be most helpful if we received it within two weeks, as we expect to submit the proposal for campus review in early October.

Please let me know if you have any questions — I would be happy to chat if necessary.

Thanks.

-Paul

_____

Paul P. Maglio - School of Engineering - University of California, Merced
Professor - Ernest and Julio Gallo Management Program
Chair - Department of Management of Complex Systems
Director - Division of Management and Information

Admin - Yolanda Kyle - ykyle@ucmerced.edu - (209) 228-3479

<DRAFT_Data Science Self-Supporting Program Proposal 20200914.pdf>

From: Paul Maglio
Sent: Wednesday, July 27, 2022 2:31 PM
To: deutsch@cs.ucsd.edu
Cc: Alvin Cha; Yolanda Kyle
Subject: Request for Comments: UC Merced Proposed Professional Master of Data Science and Analytics
Attachments: Data Science Self-Supporting Program Proposal - Clean Draft 20220726.pdf

Dear Professor Deutsch,

Nearly two years ago, we reached out for your input on our proposal for a new SSPGDP, Master of Data Science and Analytics. Our proposal underwent campus review, and we are finally ready to resubmit a revised version at the start of fall semester. But first, we are circulating our revision again to chairs of related programs across the system - per my original note below - to ask for your feedback, as required by CCGA.

If you wish to provide feedback, it would be most helpful if we received it by August 17.

Please let me know if you have any questions — I would be happy to chat if necessary.

Thanks.

-Paul

———————————————————

Paul P. Maglio - School of Engineering - University of California, Merced
Professor - Ernest and Julio Gallo Management Program
Chair - Master of Management Professional Degree Program
Director - Division of Management and Information

Admin - Yolanda Kyle - ykyle@ucmerced.edu - 209-228-3479

> On Sep 14, 2020, at 3:52 PM, Paul Maglio <pmaglio@ucmerced.edu> wrote:
>
> Dear Professor Deutsch,
>
> At UC Merced we are in the process of proposing a new graduate program leading to Master of Data Science and Analytics. In accordance with the review policy established by the systemwide Coordinating Committee of Graduate Affairs (CCGA), I am providing you, as the Chair of an existing comparable program (Master of Advanced Study in Data Science & Engineering), with a copy of the current draft of our proposal. We would be very grateful for any feedback you may wish to offer us, so that the proposal may be made as strong as possible before submission.
>
> As background, please understand that the format and contents of the proposal follow the required outline found in the CCGA Handbook, and that internal and external reviewers will later be asked to address the following four points when examining our final submission:
>
> – Quality and academic rigor of the program

– Adequacy of the size and expertise of faculty to administer the program
– Adequacy of the facilities and budgets
– Applicant pool and placement prospects for the graduates

If you wish to provide feedback, it would be most helpful if we received it within two weeks, as we expect to submit the proposal for campus review in early October.

Please let me know if you have any questions — I would be happy to chat if necessary.

Thanks.

-Paul

_____

Paul P. Maglio - School of Engineering - University of California, Merced
Professor - Ernest and Julio Gallo Management Program
Chair - Department of Management of Complex Systems
Director - Division of Management and Information

Admin - Yolanda Kyle - ykyle@ucmerced.edu - (209) 228-3479

<DRAFT_Data Science Self-Supporting Program Proposal 20200914.pdf>

Appendix J – MDSA Program Bylaws

Master of Data Science and Analytics Bylaws

Administrative Home: Proposed E & J Gallo School of Management

Revision date(s): October 15, 2020

ARTICLE I. OBJECTIVE

A. Degree offered by the program: **Master of Data Science and Analytics**

B. Discipline: **Data Science/Data Analytics**

C. Mission of the Program: The **Master of Data Science and Analytics** degree program will equip students to draw sound conclusions from data in context, using principles of statistical inference, computational processes, geographic information systems, data management strategies, domain knowledge, ethics, and probability and statistics theory. Students will learn to carry out analyses of data through the full cycle of investigative processes in scientific and managerial contexts. They will gain a deep appreciation of the human, social, and institutional structures and practices that shape technical work around computing and data, as well as an understanding of how data, data analytics, artificial intelligence, and computing permeate and shape our individual and social lives.

ARTICLE II. MEMBERSHIP

A. Criteria

Membership is open to all faculty members (including emeriti) who are core members of the following three home graduate groups: **Cognitive and Information Sciences**; **Economics;** and **Management of Innovation, Sustainability and Technology**. Faculty members who belong to more than one of these three home graduate groups maintain membership in the program as long as they maintain membership in at least one of these three graduate groups.

B. Types

All faculty members of the program are core members of the program, with all privileges of membership.

C. Voting Rights

All members are eligible to vote on program policy, bylaws, and membership petitions.

D. Review of Membership

The criteria for reviewing members are set in the bylaws of the home graduate group to which each member belongs.

E. Membership Appeal Process

Membership appeals are processed by the Graduate Group to which the faculty member belongs.

ARTICLE III. ADMINISTRATION

The administration of the Program is carried out by the Program Director and the Executive Committee. Specifically, the Program Director is responsible for the day-to-day operations of the program and serves as chair of the Executive Committee, which is responsible for policy and grievances.

ARTICLE IV. PROGRAM DIRECTOR

A. Nomination Process

The appointment of the Program Director shall be in accordance with current Graduate Council policies and Merced Academic Personnel Policies and Procedures (MAPP) 5011-5014. The Director will be appointed by the EVC/Provost for a term of 3 years upon the recommendation by the Dean of the proposed E. & J. Gallo School of Management. The Executive Committee will solicit the names of nominees from the program faculty. The names of the nominees willing to serve will then be submitted to the program faculty for comments; all comments will remain confidential. The Executive Committee will forward at least two names to the School Dean, along with comments received on the nominees.

The inaugural Director will be nominated by the graduate group chairs of Cognitive and Information Sciences, Economics, and Management of Innovation, Sustainability and Technology and will serve in the Acting capacity for one year.

The Program Director is a core member of the MDSA Program faculty, with all rights and privileges pertaining thereto.

B. Duties of the Director

The Program Director carries out the following duties:

- Oversee the progress of graduate students through the program, including satisfaction of degree requirements and advancement to candidacy, in coordination with advisors, faculty, and staff
- Represent the program faculty in all matters related to the degree program to the lead dean, the graduate dean, Graduate Council, and School Executive Committee(s)
- Determine resource needs and administer program finances (including collecting revenue, paying for instruction, returning funds to student aid,

and managing payments to the campus and graduate groups), in consultation with Executive Committee, program faculty, and lead dean

- Oversee graduate student recruitment, graduate program website, admissions, and financial aid, in consultation with Executive Committee, program faculty, lead dean, and graduate dean
- Determine graduate course offerings, including curriculum changes, in consultation with Executive Committee, program faculty, school staff, graduate groups, and departments involved in course scheduling and teaching assignments
- Determine graduate course resource needs for equipment, staff support, and other resources, in consultation with Executive Committee, program faculty, and lead dean
- Serve as Program Faculty Accreditation Organizer by overseeing annual program assessments and periodic program review, to monitor and maintain academic excellence
- Consult with Executive Committee, program faculty, and lead dean in selecting and reviewing support staff assigned to the program
- Coordinate participation of the graduate program in School and University program activities, including graduate student fellowship and award programs
- Develop and maintain a plan for promoting diversity among matriculated graduate students
- Manage and respond to program feedback and inquiries from faculty, students, staff, and reviewers

ARTICLE V. COMMITTEES

In the absence of the committees listed below (starting with B, Membership Committee), the Executive Committee will assume the duties and responsibilities for these committees.

A. Executive Committee

The Executive Committee is comprised of the Program Director, who chairs the committee, and three faculty members, one each from CIS, ECON and MIST Graduate Groups. Committee members are elected by the Program faculty in each Graduate Group, according to the voting procedures of the respective Graduate Group, for a term of 2 years. The initial Executive Committee members will have staggering terms, some serving 2 years while others serving only 1 year. Graduate Group Chairs and the Program Director will jointly decide which members would serve the 2-year term.

The Executive Committee is jointly responsible for the Program administration with the Director. The Executive Committee shall determine and implement policies for the good of the Program, including review and modification of program requirements.

The Executive Committee will meet at least once per semester. Additional meetings and executive sessions may be called by the Program Director as deemed necessary, or upon

the petition of five Program Faculty members.

B. Membership Committee

Per Article II, Membership of the Program is determined by each Graduate Group that is part of the Program. In cases where questions of the program membership cannot be resolved by the graduate group(s), the Executive Committee will function as the Membership Committee. Its final recommendation will be put to Program faculty vote at the Annual Faculty Meeting, unless an earlier meeting is called by the Executive Committee.

C. Education Policy Committee

The Education Policy Committee consists of at least three Program faculty members, one from each of CIS, ECON, and MIST. The faculty members are appointed by the Program Director, with at least two nominations solicited from each graduate group. Committee members serve a term of one year, and may be reappointed for consecutive terms. The Program Director appoints the Committee Chair.

The Education Policy Committee is responsible for reviewing the curriculum, consideration and approval of course offerings, recommendation of changes in program requirements (to be made to the Executive Committee) and making recommendations for faculty teaching assignments to Department Chairs.

D. Admissions Committee

The Admissions Committee consists of the Program Director and at least two Program faculty members appointed by the Program Director for a one-year, renewable term. The number of Committee member(s) is determined by the graduate group affiliation of the Program Director, ensuring that all three Graduate Groups are represented. A staff member whose main responsibility is recruitment of new students attends as an ex-officio member of the Admissions Committee.

Once the completed application, all supporting material, and the application fee have been received, the Admissions Committee reviews the applications and makes recommendations to accept or decline an applicant's request for admission. The recommendations are forwarded to the Dean of the proposed School of Management for final approval. This Committee may, upon Program Chair's request, assist in formulating and reviewing recruitment strategies and in organizing or attending prospective student visits to the campus.

ARTICLE VI. GRADUATE ADVISORS

Students will be advised by academic and job placement counselors who will be dedicated to serving students in this Program. The Program Director serves as the official Graduate Advisor for all students in the Program, although students' capstone projects will be advised by individual Program faculty members, or by faculty members from outside the Program, with permission from the Director.

Article VII. Meetings

A. Notification of Meetings

At least one annual meeting of the Program faculty will be held, typically at the start of Fall Semester. Other meetings may be called as frequently and for such purposes as deemed desirable by the Program Director and/or the Executive Committee. Additional meetings may be scheduled upon written petition by five or more Program faculty members, sent to the Director.

Members shall be notified of meetings at least two weeks in advance via email. Faculty away from campus may participate via teleconference or web-based conference.

B. Order of Business for Meetings

N/A

Article VIII. Quorum and Voting

Voting on program policies and other matters will done as follows:

- All matters requiring a vote must be voted on by 50% or more of the eligible members
- All matters requiring a vote, except revisions to the Bylaws, must be passed by a 50+% supporting vote of the members voting
- Revisions to the Bylaws shall require two-thirds of votes in favor to pass
- Voting may be done in a meeting of the program faculty, or by e-mail or other web-based balloting technology. The Program Director is responsible for informing the membership as to whether votes will be cast in person at a meeting or through electronic means. The default is for voting to occur by electronic means. In the case of electronic voting, ballots must remain open for at least three business days.

Article IX. Amendments

Amendments to these bylaws may be made in accordance with the quorum policy in Article IX. Program members may propose amendments by petition to the Program Director. The Director, or the Executive Committee, may ask for revisions from the faculty who submitted proposed amendments before forwarding the revisions to the membership for review and voting. Quorum, voting and passage are prescribed in Article IX. All amendments and revisions must be submitted to the Graduate Council for review and approval; changes in the bylaws will become effective upon approval by the Graduate Council.

Appendix K – Links to Curriculum Vitae for MDSA Program Faculty

COGNITIVE & INFORMATION SCIENCES		
Aguirre-Muñoz	Zenaida	https://scholar.google.com/citations?user=WD3rfYoAAAAJ&hl=en
Backer	Kristina	https://ucmerced.box.com/s/8ayh3jwclmfwudahdzrjc6yirl0fjext
Balasubramaniam	Ramesh	https://ucmerced.box.com/s/b9knwh0jf0dpojb62p1ktfelavpoqlrh
Bortfeld	Heather	https://ucmerced.box.com/s/9i2b8r80chf0dsnggddacb4i0z5qiesb
Gunn	Hanna	https://ucmerced.box.com/s/9wyqvgb2odsfcmzynntqa9bif3z79hkr
Hicks	Daniel	https://ucmerced.box.com/s/s2n8spa4l4il2sqys99kiwccqbbf1tez
Holbrook	Colin	https://ucmerced.box.com/s/cok0gdl78dblvx3dy1tzw9qwxko96ri8
Jennings	Carolyn	https://ucmerced.box.com/s/kwhk8zzss3xkmm4hg1q3nulznbdm34g8
Jennings	David	https://sites.google.com/site/platoslysis/
Kello	Christopher	https://ucmerced.box.com/s/lejxrpwq7uj6pg17smyundpfyfdgkgdq
Maglio	Paul	https://ucmerced.box.com/s/x7d9j3jufwys1utukg5h8tub5y3snstl
Marghetis	Tyler	https://ucmerced.box.com/s/jf2xvlgol22k0ae2ualkh5si6pcj8u7g
Matlock	Teenie	https://ucmerced.box.com/s/bjrgsv55ep6ms1cz91lfqao43baadjvo
Noelle	David	https://ucmerced.box.com/s/6suulxof1lk054cieudsuko1danwkgc2
Padilla	Lace	https://ucmerced.box.com/s/0mno8g4j9to3tpzf7wan0h6dcvzcohot
Ryskin	Rachel	https://ucmerced.box.com/s/eaf6as9z4zkz32m0m8a49i7qw7omwall
Smaldino	Paul	https://ucmerced.box.com/s/km5ld0tzzm4e0enkk8mb70lcvl4n8xfg
Spivey	Michael	https://ucmerced.box.com/s/gkxfobzv4fy7b343nhon29tft1zjmadh
Yoshimi	Jeffrey	https://ucmerced.box.com/s/u3odw16az49zrga190ynz1d99o58wkm0
ECONOMICS		
Amuedo-Dorantes	Catalina	https://ucmerced.box.com/s/iuvjc0fbr7y65e7err3l7ety49smz45u
Ballis	Briana	https://brianaballis.weebly.com/cv.html
Fons-Rosen	Christian	https://ucmerced.box.com/s/ovxpn78iut1wtv7ktww8fnpsgui6d74g
Gray	Rowena	https://ucmerced.box.com/s/fqzwsivtad9kw4z96dd1wzc7jdfc0w2r
Grossman	Zachary	https://ucmerced.box.com/s/t26up9orlvleh8i5jg4u2eqoxtduve4x
Hicks	Justin	https://ucmerced.box.com/s/3rkaouc8o6922yx31ptnnpnc8xzbd7d8
Innes	Robert	https://ucmerced.box.com/s/5al84t12qfa8df1pqeva4x33stfpr5zh
Johnston	Andrew	https://ucmerced.box.com/s/zqjsthdlngxnfpnfvz1qyl4vron7f6x1
Lee	Jason	https://ucmerced.box.com/s/hmjiq0q6w7dw8rim2oeqy83umyi1eas6
Sandoval-Hernandez	Jesus	https://ucmerced.box.com/s/3j4r8lf62mktco7avb8alvulb6s8swdm
Schnier	Kurt	https://ucmerced.box.com/s/zni3e8aooq2gzcwc6wrclwej3j5mna8z
Tsoulouhas	Theofanis	http://tsoulouhas.info/
Tur-Prats	Ana	https://ucmerced.box.com/s/vb38jvautrn5oo1yiqjdqq9ez32jzcqe
Wright	Greg	https://ucmerced.box.com/s/mnnb5i1hspetffgintr8bagm6sc95iji
MANAGEMENT OF INNOVATION, SUSTAINABILITY, AND TECHNOLOGY		

Abatzoglou	John	https://ucmerced.box.com/s/ev3ta7wiftz9tqg805q7ehqz5r0kde30
Bales	Roger	https://ucmerced.box.com/s/d485kary60ngdkv7crfzr11wsxknh4qb
Bhappu	Anita	https://ucmerced.box.com/s/xq4f51kfpgbxzvnal6pumkqaqnzx0v9d
Castro	Spencer	https://ucmerced.box.com/s/ih2dhhw5sig8g62iyb4lsixgjdn7r03h
Jenkins	Jeffrey	https://ucmerced.box.com/s/dybco1hqpbhg2olhiss5vmrsz9zkjzuo
Keske	Catherine	https://ucmerced.box.com/s/g5s94ybdg1tngj3h05v79obo24xdbjan
Kolden	Crystal	https://ucmerced.box.com/s/mwqq328vkkuk047pvssncz1ru1hogqrl
Lempiala	Tea	https://ucmerced.box.com/s/x0p7pxgwenun5ylw3rj3da2c389j94nj
Maglio	Paul	https://ucmerced.box.com/s/x7d9j3jufwys1utukg5h8tub5y3snstl
McBride	Russell	https://ucmerced.box.com/s/byww1udjdq3ted6j3igifru6u2b2wgqm
Osborne	Tracey	https://ucmerced.box.com/s/g4vdr24qsy1qkc8zo3xz9g510kv2pmij
Petersen	Alexander	https://ucmerced.box.com/s/rg63i89jr7s52cj4udfyqeqn92yarfsy
Westerling	LeRoy	https://tinyurl.com/ucm-lwesterling
Yeo	Lisa	https://ucmerced.box.com/s/veuyyjmvi7z6tqpar2iib1uwccczo9wn