PROPOSAL FOR AN UNDERGRADUATE PROGRAM LEADING TO A

Bachelor of Arts (B.A.) Degree in

Data Science and Analytics

Proposal Submitted By:

Alex Petersen
Associate Professor
Management of Complex Systems
School of Engineering

Proposal Development Team:

Christian Fons-Rosen, Associate Professor, Economics and Business Management
Paul Maglio, Professor, Management of Complex Systems and Cognitive and Information Sciences
Russ McBride, Assistant Professor, Management of Complex Systems
Andrew Shaver, Assistant Professor, Political Science
Jeff Yoshimi, Associate Professor, Cognitive and Information Sciences

Date: January 12, 2023

Proposed Major in Data Science & Analytics

1. Program Description and Rationale

1.1 Program Focus

We are submitting a proposal for an undergraduate *Data Science and Analytics* (DSA) major terminating in a bachelor of arts degree, to commence in Fall 2024. Coursework integrates around two skill domains: *data science* (involving critical reasoning and modeling facilitated by advanced algorithmic tools) and *data analytics* (involving the data communication, design of data services, decision-support tools and statistical inference for decision-making).[1] Consequently, the program aligns with two CIP codes: *30.7001 - Data Science, General* and *30.7101 – Data Analytics, General*, both of which belong to the "multi/interdisciplinary studies" category[2]. Accordingly, four existing departments – Economics and Business Management (EBM), Cognitive and Information Sciences (CIS), Management of Complex Systems (MCS), and Political Science (POLISCI) – will co-deliver and co-govern the program.

The proposed DSA major supports the campus' longstanding commitment to developing the broad area of *Computational Science and Data Analytics*, as exemplified in a 2016 multi-disciplinary cluster-hire of early-career faculty from this domain.[3] As an attractive new major, offered by existing faculty, and aligned with growing workforce demand, we strongly believe the DSA program will contribute to campus enrollment-growth objectives.[4] The DSA curriculum accommodates students switching majors from the CSE program after the first and second years with no anticipated impact on normative time-to-degree; it also readily accommodates 3rd year CA community college transfers. The DSA program also supports interdisciplinary educational pathways in support of the campus' stated mission to "foster and encourage cross-disciplinary inquiry and discovery". Faculty from the four departments have demonstrable research and teaching experience in the domains of ethics, data and information science, inferential methods, and systems thinking – and as such, are prepared to educate students in applied data science,[5] including best practices of data collection, integration, visualization, inferential analysis and decision-making.[6]

Although the proposal was developed jointly by these four departments, which span two UC Merced schools (Social Sciences Humanities and Arts; and Engineering), it will be launched and initially supported by the School of Engineering (i.e., the source of advising and other

[1] We strongly endorse the development, diversification and inclusion of various undergraduate and graduate data-science programs at UC Merced, and anticipate various ways that multiple programs can contribute to programmatic breadth, resiliency and dynamism. As such, we are supportive of other programs also using any combination or derivative of "Data Science" to identify their program.
[2] The proposed DSA program will summarily be classified by the "Data Science, General" CIP code *30.7001*.
[3] Update on the Strategic Academic Focusing Initiative, UC Merced Panorama, May 20, 2015.; Key Themes to Drive Future Faculty Recruitment Campus's First Set of Cluster Hires Comes Aboard. UC Merced Panorama, August 28 2015; UC Merced Panorama, Sept. 7 2016.
[4] See Table 5 which illustrates the straightforward course pathway for a UC Merced CSE major who transfers after completing that major's standard first-year courses; and see Table 6 for a hypothetical transfer pathway for a incoming 3rd-year Merced Community College student.
[5] National Academies of Sciences, Engineering, and Medicine. (2018). Envisioning the data science discipline: the undergraduate perspective: interim report. (2018). Data science for undergraduates: Opportunities and options. National Academies Press.
[6] Börner, K., Bueckle, A., & Ginda, M. (2019). Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. PNAS, 116(6), 1857-1864.

undergraduate education staff support). The undergraduate program will migrate to the E & J Gallo School of Management (comprised of the same four departments) once this new school is approved and launched, possibly in 2024 or with a 1-year lag. For clarity, the DSA proposal and the Gallo School proposals are not contingent, and so they should be considered as separate proposals even though they derive from a common set of departments and faculty.

1.2 Program Rationale

1.2.1 Workforce and programmatic demand

Both academia and industry have strong demand for computationally adept analysts capable of extracting value from data by integrating into teams tasked with large-scale workflows[7]. Stages of the workflow include collecting data from a myriad of existing databases and other less structured data sources; integrating these data into common data formats; analyzing and modeling multi-variable relationships; and communicating via graphical visualization or by designing interactive web platforms that strategically facilitate data-driven decision making. In addition to technical know-how, professional analysists should also be grounded in the data ethics[8] associated with collection, ownership, privacy and systemic bias that are of prevailing concern regarding human-data based research and industrial data platforms.

California is home to numerous DSA-oriented industries and research universities. To address this competitive workforce demand for DSA skills in addition to an emerging digital skills gap, starting as early as K-12 education and well-documented in the present data economy,[9] an increasing number of universities offer undergraduate data science programs of varying intensity and modality (online, in-person)[10]. It is not uncommon that a university offers multiple programs delivered by different sets of faculty to accommodate varying approaches and levels of technical rigor in the treatment of data science methodologies.[11] As such, we anticipate robust initial enrollment and sustainable growth for this proposed major.

1.2.2 Novelty in support of campus' cross-disciplinary mission

An attractive component of the major is the diversity of faculty and coursework, which provide an interdisciplinary blend of perspectives (e.g. engineering-inspired design thinking, systems thinking) and methods (survey instruments, econometric inference) that are crucial for addressing real-world problems. Moreover, the core problem domains of the four departments organically integrate *People, Planet and Profit* (PPP) sustainability dimensions[12] are aligned

---

[7] Börner, K., et al. (2018). Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy. Proceedings of the National Academy of Sciences, 115(50), 12630-12637.

[8] Oliver, C. and McNeil, T. Data science education lacks a much-needed focus on ethics. The Conversation, 9/2021.

[9] In addition to [6] see also: Cappelli, P., 2012. Why good people can't get jobs: The skills gap and what companies can do about it. Wharton Digital Press; Jackman, J. A., et al. (2021). Addressing the digital skills gap for future education. Nature Human Behavior.

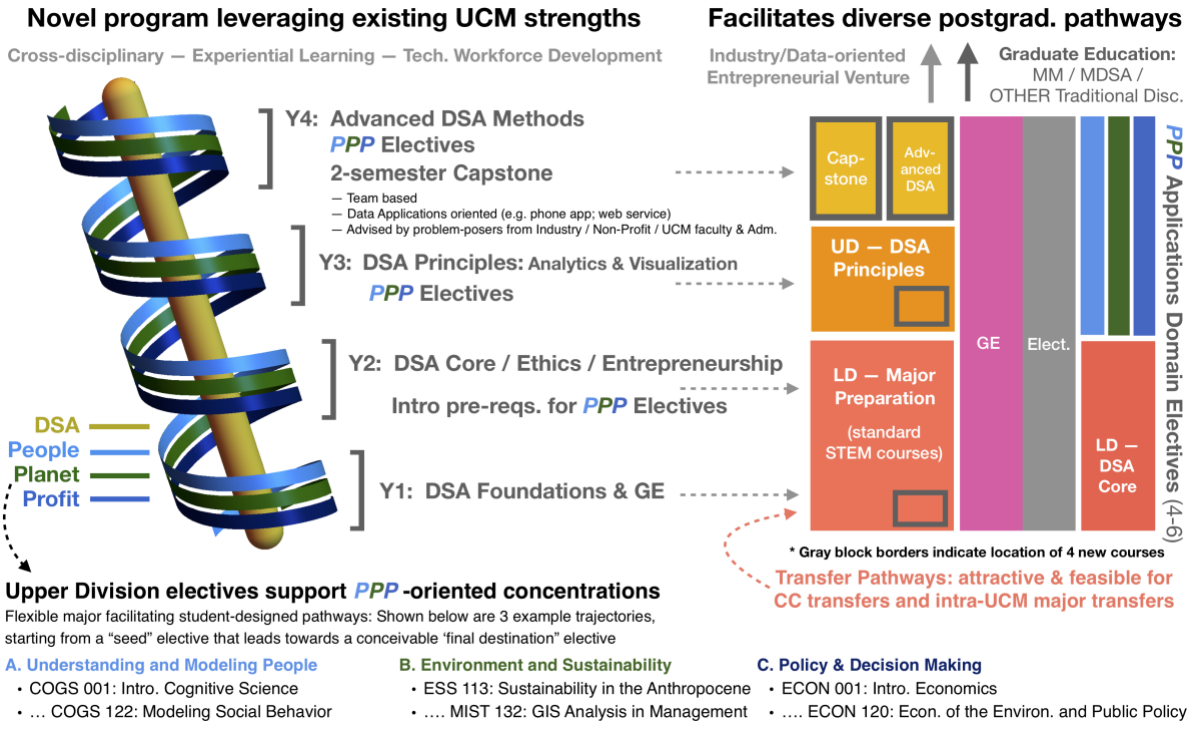
[10] Two lists of undergraduate data science programs, one complied in June 2019[10] and a recent US News 2021 ranking,[10] indicate a variety of DSA programs ranging from certificates, to minors, degree concentrations, and standalone bachelor degrees. The California Alliance for Data Science Education lists regional data science major offerings.

[11] For example, the University of Michigan launched two "Data Science" programs in 2016, thereby demonstrating the appeal for both deep and broad cross-disciplinary approaches to delivering DSA education at the undergraduate level. Another example of a public university offering multiple data science pathways to appeal to different student backgrounds and needs, is Penn State University, which offers *three* Data Science bachelors programs, one in each of three schools.

[12] Savitz, A. (2013). The triple bottom line: how today's best-run companies are achieving economic, social and environmental success-and how you can too. John Wiley & Sons.

with existing programs (e.g. the Master of Management) as well as the emphasis on sustainability in the campus' mission statement.

Figure 1 shows how flexible application-domain electives that leverage existing courses and faculty strengths in three broad *PPP* areas: *Understanding and Modeling People*; *Environment and Sustainability*; *Policy and Decision Making*. Hence, graduates equipped with a T-shaped skillset integrating both technical know-how (data science) and relevant soft skills (data analytics) will be competitive for a wide range of career pathways, including competency in the analytic toolset fundamental to graduate education in a wide range of STEM and interdisciplinary programs.



<u>Figure 1.</u> DSA educational framework at the nexus of People, Planet and Profit (PPP) sustainability perspectives, thereby integrating around the campus' mission and existing faculty strengths. (Left) *PPP* electives provide context for data-driven education, inquiry and real-world applications. Students in Year 1 (Y1) complete foundational methods courses and GE requirements, similar to other undergraduate STEM degree programs, thereby facilitating efficient transfer in and out of the proposed program to other STEM majors; Y2 introduces computational, as well as introductory ethics and entrepreneurship to guide translating ideas into practice; Y3 introduces core DSA principals, and is where students begin to broaden application-domain knowledge by way of select PPP electives; Y4 introduces advanced methods relevant to job market and graduate school demand, and orients students towards cross-boundary DSA applications, in particular via an integrative team-based Capstone project (2-semester) where a real-world problem is posed by an industry advisor or local UCM faculty or even a campus administrative unit (e.g. the Library or Facilities Management). (Right) Gantt chart shows curriculum comprised mostly of existing courses. Only 4 new courses (DSA 002; 101 or 102; 103; 120/121) would be needed to launch the DSA program. (Bottom) Students identify at least three electives that enrich their understanding of the endless frontier of DSA applications.

1.2.3 Overall Value Proposition

The proposed program does not require any new faculty lines, staff or specialized space, facilities or equipment, and requires just four new 4-unit courses: DSA 002, DSA 101 or DSA 102, DSA 103 and DSA 120/121. The resource requirements are also relatively small owing to the low material cost of delivery (leveraging students' personal laptops as primary computing device for accessing open-source coding platforms and free data repositories, which aligns with new School of Engineering policy that is also being considered for the campus; see Sec. 4.4 for more elaboration on instructional computing resources).

Given the proliferation of online certificate programs in data science methods, another important value proposition and comparative advantage of the proposed program is the experiential learning gained through an integrative senior year capstone (i.e. a problem- and team-oriented senior research project), where teams will apply DSA methods to solve real-world challenges identified by campus research advisors and industry partners.

1.3 Contribution to undergraduate education at UCM

The proposed undergraduate DSA major will horizontally integrate soft data-communication skills and technical know-how that vertically integrates with a deep understanding of real-world *People, Planet and Profit* (PPP) application domains. More specifically, the horizontal integration is comprised of two complementary skillsets - *data science* (oriented around measurement, critical reasoning and modeling) and *data analytics* (involving data-driven decision-making, decision-support tool design, and real-world data services deployment).

This toolkit will allow students to pose and address questions in very broad fields of study spanned by the program faculty, including human behavior, agriculture, environment, law, innovation, management, economics and political science. Another distinguishing feature of the proposed curriculum is the integration of data ethics, data-services entrepreneurship, and problem-driven data challenges encountered in areas like climate change or social justice. The resulting T-shaped skillset emerges organically from the convergence of the four departments delivering the program, and is befitting of other transdisciplinary data science and data analytics programs.[13,14]

The curriculum structure is rather flexible by way of electives, which supports completion within the normative 4-year graduation plan. Lower and upper-division courses satisfy nearly all of the GE requirements.

1.3.1 Alignment with undergraduate DSA education across the UC

Undergraduate data science programs are no longer found just at the margins of elite private and flagship public universities, but are becoming standard. By way of example, recent data compiled by the *California Alliance for Data Science Education* shows 65 California institutions developing or having already established programs in this direction.

[13] Why Colleges Are Offering Data Science Programs, US News, 2020.
[14] In particular, MIT offers a novel undergraduate major in *Computer Science, Economics and Data Science*: see Two Sciences tie the knot, MIT News, 2017; Degree description.

Importantly, the proposed DSA program will bring UC Merced into alignment with the emergence of undergraduate data-science educational programs across the UC system in the last decade, in particular at UCB, UCI, UCR, UCSD, and UCSB. The most similar data science program within the UC system is the Data Science degree offered by the UC Berkeley (UCB) College of Letters & Science[15]. See section 1.7 for further elaboration.

1.3.2 Alignment with PPP graduate education at UC Merced

The triple-bottom-line (People, Profit, Planet) framework for the application domain electives aligns with the existing Master of Management (M. M.) program offered by faculty from the Management of Complex Systems department. The M. M. program would be a natural 5th year management education option, one that complements the technical and application-domain education provided by the DSA major. Consequently, graduates of the DSA major would be ideal candidates for the M. M. program.

1.3.3 Alignment with existing undergraduate education delivered by the four departments

The proposed curriculum incorporates more than 20 existing data-oriented courses offered across the four departments – see *Appendix I* for course descriptions. It also complements the existing Management Analytics and Decision-Making (MAD) minor offered by faculty from the Management of Complex Systems department. Graduates of the DSA program would not be able to declare this minor, as the MAD minor is satisfied by core requirements of the proposed DSA major. However, the MAD minor would be a possible pathway for students transferring into the DSA major.

1.3.4 Complement to existing computing- and quantitative-intensive majors: CSE and Applied Math

The proposed program complements existing technical computation-oriented programs on campus (e.g., Computer Science and Engineering (CSE), Applied Mathematics) as well as planned future programs (e.g. a Data Science major under development within the School of Natural Sciences), which tend to specialize in the development, deployment and optimization of data-oriented algorithms and their applications in both traditional and interdisciplinary STEM domains. This is in contrast to the proposed DSA program, which introduces the methodology for using technical methods, but does not focus in their development and optimization; and instead focuses on their deployment in econometric analysis, decision-support tools and entrepreneurial ventures. As such, we envision the proposed DSA program contributing to the emergence of a campus-wide data science culture comprised of complementary cross-supporting educational and extracurricular programs.

1.4 Job market

The program will prepare students for both existing and emergent professions in data science and data analytics in the private sector, international organizations, and governmental agencies, both in the Central Valley and across the globe. The skills learned in this major can help students in their goal of becoming data managers, entrepreneurs, evaluating colleague's data analytics work at their job, academics in data-related social and life science fields, school and college educators, or data journalists, among many others.

[15] https://data.berkeley.edu/academics/data-science-undergraduate-studies/data-science-major ; https://data.berkeley.edu/degrees/domain-emphasis

1.4.1 Job types and skills required

UC Merced's Office of Institutional Research & Decision Support (IRDS) compiled a job market analysis based upon the CIP codes: *30.7001 - Data Science, General* and *30.7101 – Data Analytics, General*, which indicates a flourishing market for these skills, with upwards of half a million jobs in this area in California alone by 2030.[16] Accordingly, we anticipate that graduates of this program will address Central Valley workforce development needs – whether by seeding an information-based economy through the development of a local supply of data scientists and related start-ups, or by developing remote sensing and other techno-informatic management capabilities critical to *Precision Ag.*[17] As a reflection of the campus' diverse student body, graduates will provide underappreciated perspective[18] to the nation's increasingly data-intensive workforce.[19]

Specific job types identified in external assessment are *Data/Data Mining Analyst, Data Scientist, Statistician, Business Intelligence Analyst/Architect/Developer, Market Research Analyst, Financial Quantitative Analyst,* and *Research Manager* (see the Job Market Assessment report in the *Appendix II - IRDS*).

Particular skills associated within Data Science and Data Analytics job postings in California that this program will provide are *Communication, Teamwork,* computing fundamentals (*Python/SQL/Git*), *Data Visualization*, *Project Management* and *Business Intelligence, Economics* – i.e., critical thinking and communicating with data (see the Job Market Assessment report in the *Appendix II - IRDS*).

1.4.2 Professional DSA Industry prospects

A 2017 report[20] identified DSA as one of the most rapidly expanding pathways leading to careers in finance, professional services, IT and other sectors. While about 40% of these jobs would require an advanced degree, such as a master's in data science and analytics, the remainder could be filled by graduates with a bachelor's degree in DSA. 98% of hires have a minimum of a bachelor's degree. In addition to strong historical demand featuring competitive starting salaries, DSA-oriented industries are also anticipated to be resilient to present economic and labor market challenges.[21]

---

[16] See *Appendix II - IRDS* where we include the relevant IRDS reports using Burning Glass assessment tools, based upon the Classification of Instructional Programs (CIP) queries "30.7001 - Data Science, General" and "30.7101 – Data Analytics, General". Notably, both of these CIP categories are found in top-level category 30 - "Multi/Interdisciplinary Studies", which the CIP defines as "Instructional programs that derive from two or more distinct programs to provide a cross-cutting focus on a subject concentration that is not subsumed under a single discipline or occupational field.", which is consistent with and provides support for this multi-disciplinary program design.

[17] National Academies of Sciences, Engineering, and Medicine, 2018. Improving Data Collection and Measurement of Complex Farms. National Academies Press.

[18] National Academies of Sciences, Engineering, and Medicine, 2018. Minority serving institutions: America's underutilized resource for strengthening the STEM workforce. National Academies Press.

[19] Information technology and the US Workforce: Where are we and where do we go from here?. National Academies Press, 2017; Promising practices for strengthening the regional STEM workforce development ecosystem. National Academies Press, 2016; Strengthening Data Science Methods for Department of Defense Personnel and Readiness Missions. National Academies Press, 2017; Future US workforce for geospatial intelligence. National Academies Press, 2013.

[20] The Quant Crunch: How the demand for data science skills is disrupting the job market, Burning Glass Technologies, 2017.

[21] Say goodbye to six-figure starting salaries – with these exceptions, CNBC, 2020.

According the LinkedIn's 2020 Emerging Jobs Report, DSA jobs have been growing at roughly 37% annually and are part of the fastest three growing jobs categories (which also includes behavioral health professionals and engineers).[22] Data scientists earn an average of $121,189 annually. Interestingly, three of the five cities with the highest average pay for data scientists are all in California, with the nearby Bay Area leading the salary ranking[23]:

1. San Francisco average DSA salary, $170k, adjusted for cost of living, $128k
2. Boston average DSA salary, $135k, adjusted for cost of living, $120k
3. San Diego average DSA salary, $142k, adjusted for cost of living, $119k
4. Chicago average DSA salary, $116k, adjusted for cost of living, $107k
5. Los Angeles average DSA salary, $129k, adjusted for cost of living, $106k

Note that averages that tend to be skewed by extreme values. Also note that DSA students will be trained in entrepreneurial and management fundamentals that equip students to potentially shift into leadership positions that tend to have relatively high.

Figure 2 shows demand for data scientists increasing exponentially over time, as the number of professionals entering the labor market who refer to themselves on LinkedIn as "data scientists" increases with each year.[24] More recently, a LinkedIn report from 2018 identified 151,000 unfilled data science jobs.[25] In a parallel effort yielding strong statistical support, UC Merced's Office of Institutional Research & Decision Support (IRDS) also prepared an analysis of job market size and growth using Burning Glass assessment tools: to summarize they report an estimated "394,578 Data Science or Data Analytics jobs in California in 2015, increasing to 407,782 Data Science or Data Analytics jobs in 2020 and predicted to grow to 488,891 Data Science or Data Analytics jobs in 2030, which is high relative to growth in other fields. The Burning Glass assessment found that median California salary for graduates with a bachelor's degree in Data Analytics would be $103,000"; for more details, see *Appendix II - IRDS*.

Figure 3 indicates that the greatest demand among the various subfields within the broad field of data science appears to be general DSA analytic skills and the ability to handle big data, the top of the skill list below (as opposed to A.I., robotics, and other sub-fields); another prominent skills vacancy is business analysis, the 6th skill below.[26]

[22] Berger, G. (2020, January). LinkedIn 2020 Emerging Jobs Report. https://business.linkedin.com/content/dam/me/business/en-us/talent-solutions/emerging-jobs-report/Emerging_Jobs_Report_U.S._FINAL.pdf Retrieved from URL.
[23] Rayome, A. DeNisco. (Aug 12, 2019). Data scientists earn the highest salaries in these 5 cities. https://www.techrepublic.com/article/data-scientists-earn-the-highest-salary-in-these-5-cities/ Retrieved from URL.
[24] Gilpress (Nov 8, 2015). Top skills and backgrounds of data scientists on LinkedIn. https://whatsthebigdata.com/2015/11/08/top-skills-and-backgrounds-of-data-scientists-on-linkedin/
Retrieved from URL.
[25] LinkedIn Workforce Report | United States | August 2018. Retrieved from URL.
[26] Woodie, A. What's driving data science hiring in 2019? https://www.datanami.com/2019/01/30/whats-driving-data-science-hiring-in-2019/ Retrieved from URL.

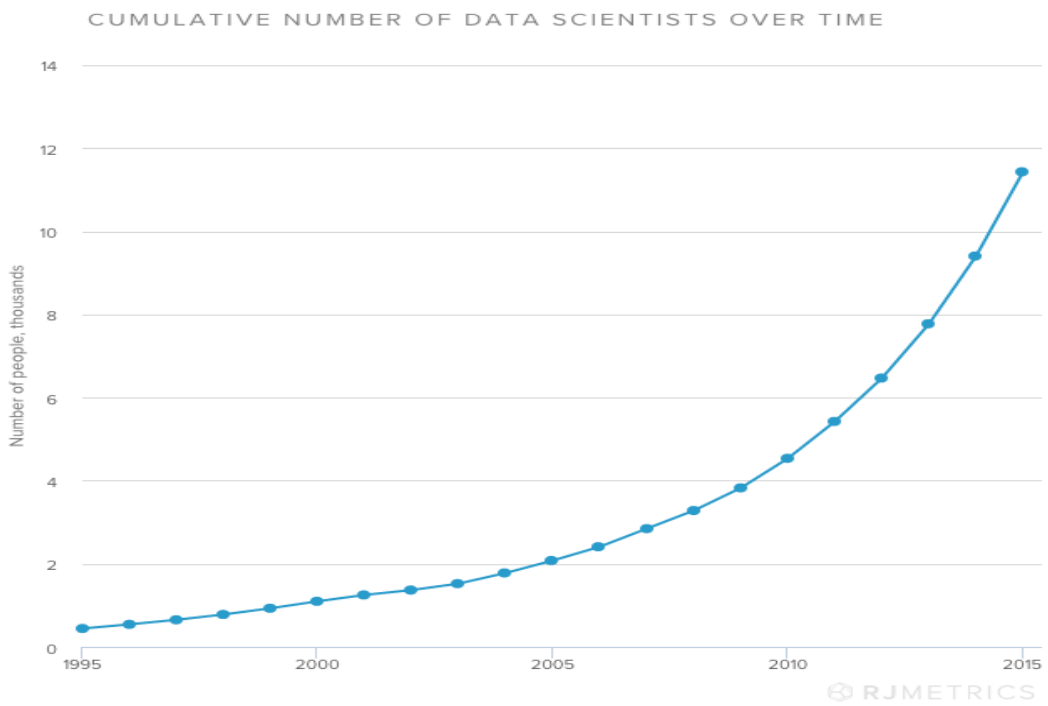


Figure 2. **Accelerating demand for DSA professionals. Image sourced from whatsthebigdata.com.**[19]

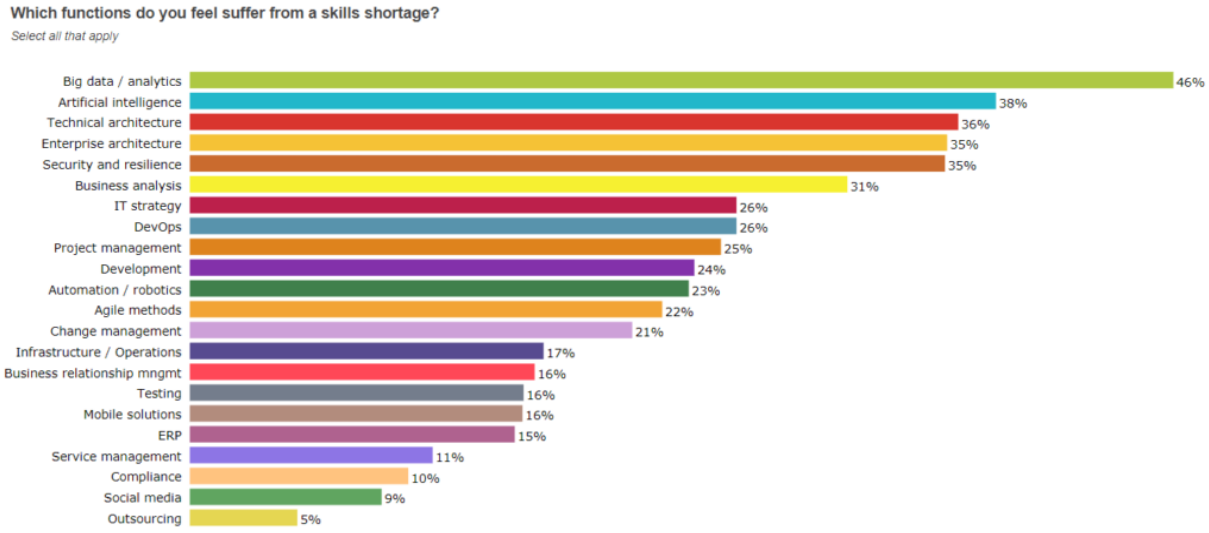


Figure 3. **Demand by DSA skillset. Image sourced from datanami.com .**[21]

The demand for DSA-trained students appears to be accelerating. The same trends are occurring in Europe. According to a report from ITPro,

"Simply put, there aren't enough data scientists to go around – the demand for data analysis has grown exponentially over the last few years, and there aren't enough people being trained to meet the demand. This growing gap between demand and available talent has meant that almost half of all European companies are thought to be struggling to fill their data scientist positions".

Consequently, 47% of organizations are struggling to fill their data science positions.[27] According to the Senior Vice President of Global Data Product Management at PepsiCo, Tony Ralph (personal communication), *"we simply can't hire data scientists fast enough"*. (Ralph taught himself data science and worked his way through Silicon Valley until he was earning $1.2M annually running a data science ad analytics team for Walmart).

Hence, the job market demand is booming in the corporate sectors, typically used to improve software, better target advertising efforts, or better analyze customer use, the demands for those with DSA training are extending well beyond typical use cases and into the humanities fields, non-profit organizations, environmental and sustainability work, and policy decision-making, among other areas (c.f., Townsend, 2021, e.g.).[28]

1.4.3 Graduate education and professional school prospects

The interdisciplinary composition of the curriculum develops a broad T-shaped knowledge and skillset combination by way of core technical coursework combined with social, economic, and environmental application/contexts via PPP electives. This composition is complemented by a highly multi-disciplinary faculty. As such, we believe that graduates of the program would be well-prepared to continue onto interdisciplinary graduate programs, such as the Environmental Systems and Management of Complex Systems PhD programs, as well as the 1-year Masters of Management, which are all presently offered at UC Merced.

1.5 Expected student enrollment

The undergraduate Data Science major is a relatively new paradigm. Hence, there are limited available data sources to draw upon regarding long-term enrollment growth at public universities of similar size and mission as UC Merced. This consideration highlights the opportunity for UC Merced to establish a prominent set of programs in this domain that will enable a early-mover advantage in the Central Valley, that could then be positioned to compete in the broader regional ecosystem of interdisciplinary and applied data science education.

1.5.1 University of California

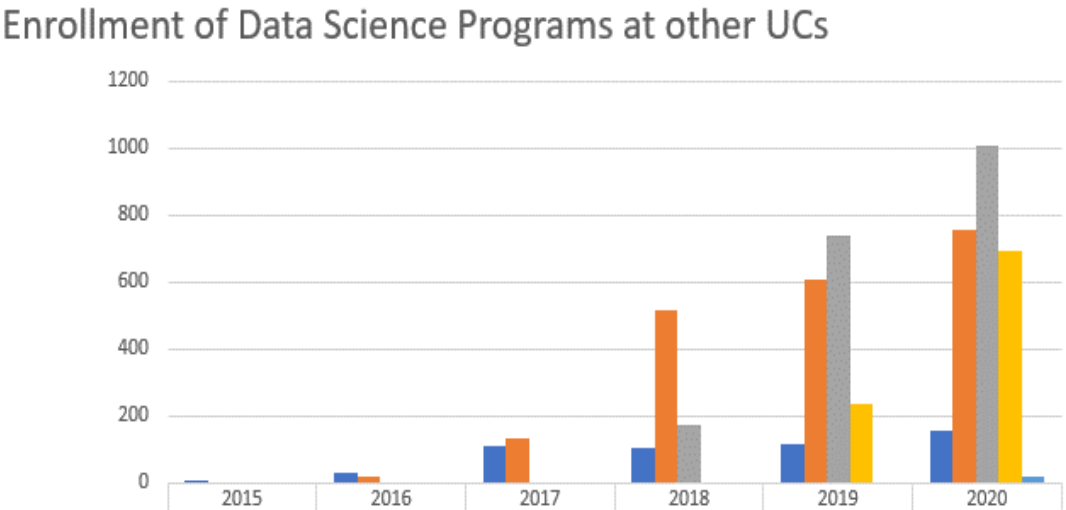
Figure 4 shows enrollment data for five UC campuses offering undergraduate degrees in Data Science in AY 2020-21. UC Berkeley offers a Data Science major through the College of Letters & Science where enrollment has exceeded one thousand students in just the first three years from fall 2018 to fall 2020 (data link). Notably, the UCB program is comprised of 28 Domain

[27] ITPro Team. (2019, Sept 17). Data scientist jobs: Where does the big data talent gap lie? https://www.itpro.co.uk/careers/28929/data-scientist-jobs-where-does-the-big-data-talent-gap-lie Retrieved from URL.
[28] Townsend, K. (2021, June 29). Why, how, and what of data science for social impact. https://data.org/news/why-how-and-what-of-data-science-for-social-impact/ Retrieved from URL.

Emphases (or tracks), that are oriented around three pillars: Computing, Data Science, and Society. By comparison, the proposed UCM DSA program is designed around three tracks that leverage the campus' existing strengths in applied sustainability themes (PPP), and so in this regard, is less focused around a technical computing and algorithm design curriculum common to other computing-oriented programs.

UC Irvine Donald Bren School of Information and Computer Sciences offers a Data Science B.S., which commenced in AY 2015-16; for fall 2020 UCI recorded 157 student-major enrollments (data link). The UC San Diego Halıcıoğlu Data Science Institute launched an undergraduate major in AY 2017, which graduated its first undergraduate cohort of 49 students in AY 2019-20 (data link).

At UC Santa Barbara, the Department of Statistics and Applied Probability oversees both a B.A. and a B.S in Statistics and Data Science with the difference being that the B.A. requires less upper-division units and allows flexibility in courses for the major. The program was first offered fall 2019 with an enrollment size of 237 students (data link). UC Riverside is the most recent campus to offer a B.S. in Data Science beginning in the fall 2020 term. The degree program is a collaboration between two colleges, Bourns College of Engineering and College of Natural and Agricultural Sciences, and had an initial enrollment of 18 students (data link).



	2015	2016	2017	2018	2019	2020
UCI: Data Science, B.S.	4	29	111	106	113	157
UCSD: Data Science, B.S.		16	131	517	609	759
UCB: Data Science, B.A.				171	739	1007
UCSB: Statistics and Data Science, B.A./B.S.					237	696
UCR: Data Science, B.S.						18

Figure 4. Enrollment of Data Science Programs at University of California Locations. Data are obtained from individual Planning Analysis and Institutional Research web page for UCI, UCSD, UCB, UCSB, and UCR .

1.5.2 Other Public Universities

Figure 5 shows total academic-year course enrollment numbers from another public university, University of Michigan, which is home to two "Data Science" majors (both of the same name). These two majors are offered by two different colleges – the College of Literature, Science and the Arts (LSA) and the College of Engineering. These programs thus offer a complementary perspective on enrollment growth, despite having a common framing and name. Since launching in 2016, annual counts of student course enrollments shows that the major offered by the U. Michigan LSA program grew faster than its Engineering counterpart, and also features greater gender balance. The approach Michigan took also illustrates that in a field as broad as data science there is room, and indeed demand, for multiple perspectives and approaches to it. Another examples of multi-program coexistence include Penn State University, where three different schools offer three different bachelors programs in Data Science (Link).

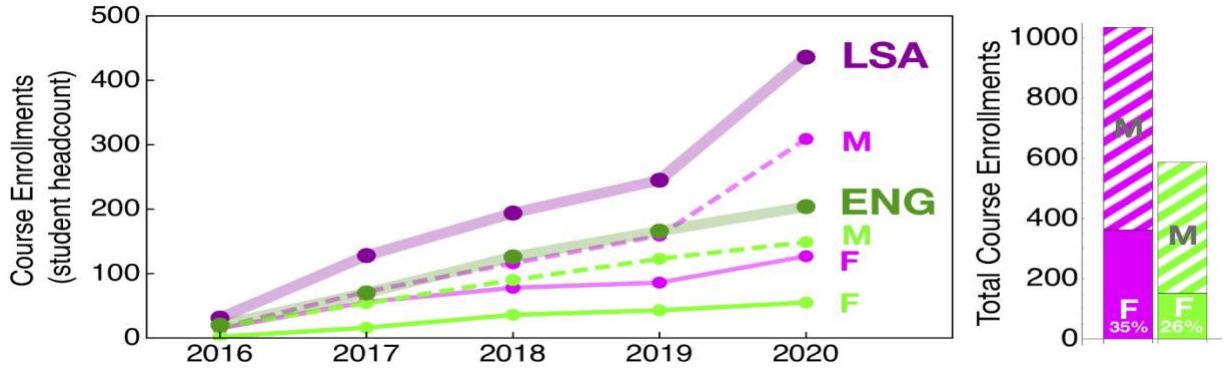


Figure 5. **Counts of student-course enrollments shown as Academic Year totals** – for two undergraduate "*Data Science"* programs offered by the University of Michigan (official records from the UM Office of the Registrar – "Field of Study" data link). School of Engineering program in Green; School of Literature, Science and the Arts (LSA) program in Magenta. (Left) Enrollment growth over time: Thin solid - Female; Thin dashed - Male; Thick solid - Total (F+M). Counts are total number of students enrolled in data science courses tallied across Fall, Winter, Spring and Summer terms; hence, dividing by 3 approximates the total number of individual students enrolled in the major. (Right) Total enrollment by program and gender group – the LSA program has higher enrollment and higher gender parity.

1.5.3 Anticipated UC Merced Enrollment

Based upon the data from UC campuses, as well as from other public universities, there is demonstrated opportunity for expanding enrollment in the broad domain of DSA. To project anticipated enrollment at UC Merced specifically, we consulted with the campus' Academic Planning and Assessment Support Center and Institutional Research & Decision Support Center to obtain an external projection of initial enrollment and growth. The full report is provided in the Appendix Documents folder, document titled "**Data Science Student Projections 12.6.2022.docx**", and the most relevant results are summarized below. The analysis starts by using 2021 UC referral pool data to provide a first estimate of present demand for data science education among UC applicants. In this regard, 724 applicants in the referral pool (representing 2% of the total) selected "data science" in their UC application, which demonstrates considerable unmet demand across the UC system.

The analysis then estimates combined enrollment at UC Merced for students selecting a data science major. Given that there are at least two planned data science programs, the following numbers approximate total enrollment in Data Science programs at UC Merced, with some portion selecting the proposed DSA major. Given the distinct orientations of the DSA major and the planned School of Natural Science major, it is reasonable as a first approximation that the demand would be split roughly equally across the available majors, and so dividing estimates by 2 provides an estimate for the DSA major enrollment, specifically.

With this in mind, the combined enrollment estimates based upon SSHA student yield rates project 2030 data science enrollment at UC Merced to be between 290 ("Scenario 3") and 605 students ("Scenario 4"). These estimates incorporate empirical growth rates of new majors and retention rates for SSHA majors. Copied below are the annual projections for the conservative Scenario 3, which suggests a total enrollment of 187 students by Year 5 of the program, growing from an initial cohort of 23 students in Fall 2024. See the full report Table 9 for the less conservative estimate, which indicates a total enrollment of 386 students by Year 5 of the program, growing from an initial cohort of 46 students in Fall 2024.

One additional factor that support these estimates being on the conservative side is they do not account for intra-UC Merced transfers from the CSE major as new data-science educational pathways arise.

Table 1 New Student and Retention Projections - 290 Data Science majors by 2030

(Scenario 3 results copied from Table 7 in the full report - Data Science Student Projections 12.6.2022.docx)

	Fall 2024	Fall 2025	Fall 2026	Fall 2027	Fall 2028	Fall 2029	Fall 2030
Frosh Growth Rates		35%	43%	34%	31%	17%	10%
New Frosh	23	31	44	59	78	91	100
Transfer Growth Rates				0%	65%	96%	4%
Transfer			3	3	5	10	10
Frosh Retention							
1st Year	81%	19	25	36	48	63	74
2nd Year	68%		16	21	30	40	53
3rd Year	64%			15	20	28	38
4th Year	11%				3	3	5
Transfer Retention							
1st Year	88%			3	3	4	9
2nd Year	33%				1	1	2
Data Science Projected Enrollments							
	Fall 2024	Fall 2025	Fall 2026	Fall 2027	Fall 2028	Fall 2029	Fall 2030
Total Enrollment	23	50	88	137	187	242	290

1.6 Program elements

The proposed DSA program is an interdisciplinary program that integrates social science (cognitive, economic, complexity, team science) and design-thinking (engineering, management) perspectives with data-oriented capabilities, with the objective of providing students a holistic training in the sources, applications, and ramifications of data.

The design of the proposed DSA curriculum is similar to most undergraduate B. Sc. majors. The first two years of the curriculum are comprised of three components: (a) general education; (b) major preparation: STEM fundamentals (e.g. calculus, introductory computing) and social science fundamentals (e.g. introduction to economics and entrepreneurship); (c) and core fundamentals (e.g. fundamentals of data analytics, statistical inference, entrepreneurship).

The third year develops DSA principals (e.g. econometrics, data visualization, data ethics) combined with flexible electives that allow students to explore relevant PPP application domains. And the fourth year introduces advanced methods of DSA (e.g. interactive visualization, machine learning and natural language processing) against the backdrop of a culminating experience applying coursework in a year-long team-based capstone project.

The program does not involve travel, field work, exposition, or any other element common to more non-traditional programs.

1.7 Overlaps and complements to existing undergraduate programs

The three most-similar majors at UCM are: Economics, Computer Science and Engineering (CSE) and Applied Mathematics major with Computational and Data Sciences Emphasis. However, these majors require students to master theoretical foundations including, but not limited to, advanced mathematical representation and modeling, high performance computing, and algorithm design.

The proposed DSA curriculum does not employ advanced mathematical methods, such as derivation and optimization, to introduce relevant quantitative foundations. In this regard, it is similar to other B. Sc. Majors such as the Biological Sciences major, which requires just two lower-division courses (MATH 011, 012) that are service courses provided to all STEM and social science majors across the university.

Below we elaborate on how the proposed DSA program overlaps and complements the two existing programs that provide data-oriented education in much more theoretical and quantitative depth, namely CSE and Applied Math. In summary, despite having different educational missions, we nevertheless anticipate strong complementarities among these various data-oriented programs that will broaden extracurricular student opportunities, e.g. by way of 'hackathon' events that stimulate and integrate the campus' broader data science community.

1.7.1 Relation to the Computer Science and Engineering Major

The Computer Science and Engineering (CSE) major is a high-demand major at UCM, comprised of a rigorous curriculum in core computer programing, algorithms, database, architecture, and operating systems. After multiple direct consultations with CSE faculty and

chair, we believe the proposed DSA major will serve as a demand shunt to relieve present over-capacity demand for the CSE major, by providing an alternative soft-skills pathway for students seeking a career in broad domain of data science and analytics. As such, by design there is little overlap in the programmatic course-work and mission of the proposed DSA major and the CSE major, aside from the computer serving as the vehicle.

Regarding the overwhelming demand for CSE courses at UC Merced and in general, we were also mindful to not exacerbate this issue, such that the proposed DSA program internalizes the introductory (DSA 001 and 002) and other intermediate (MIST 130, MIST 134) and advanced computation-oriented courses (DSA 101, 102).

1.7.2 Relation to the Applied Math. Sciences Major - Computational and Data Sciences (CDS) Emphasis Track

The Applied Math CDS emphasis track requires students to select 3 out of 5 specific courses (12 units) that introduce advanced computational methods associated with differential equations, stochastic (random number) modeling, advanced statistics, linear algebra and optimization. These courses are all upper division, and enrollment impact (in either positive or negative sense) is anticipated to be small, as the upper division MATH courses cover advanced topics that only a small number of select DSA students would likely seek to take. The proposed DSA program is not planning to introduce new courses in any of these areas, but would certainly be open to enrolling students from this and other related majors in DSA courses.

The distinction between the Applied Math CDS track and the proposed DSA Major are thus illustrated by comparing the respective program PLOs. The proposed DSA PLOs are outlined in Section 2.3 (link); for convenience, we provide the Applied Math PLOs below (program description link):

1. Solve mathematical problems using analytical methods.
2. Solve mathematical problems using computational methods.
3. Recognize the relationships between different areas of mathematics and the connections between mathematics and other disciplines.
4. Give clear and organized written and verbal explanations of mathematical ideas to a variety of audiences.
5. Model real-world problems mathematically and analyze those models using their mastery of the core concepts.

1.7.3 Relation to the planned UCM School of Natural Sciences data science major

The UCM School of Natural Sciences (SNS) is developing a data science major that emphasizes technical dimensions of mathematics, statistics, optimization and modeling applied to the domains of applied mathematics, biology and environmental systems sciences, as encompassed by the corresponding existing majors with the same names. See *Appendix III* for a side-by-side comparison of the course requirements and electives.

As described in their program description, we see the SNS program appealing to a distinct type of student by way of a more technically oriented curriculum than is being proposed here. Because the lower-division and upper-division course requirements are largely distinct, we do

not anticipate students waffling between the programs to the degree that, for example, School of Engineering students might between the engineering majors. This important distinction is also evident in the application domain areas between the two programs, which only intersect in the domain of environment and sustainability, which are shared campus pillars. Hence, the clear distinction will ensure that students, with the guidance of their academic advisor, can identify the appropriate program curriculum and application domains that suits their career interest and propensity for mathematical foundations of data science. One guiding factor supporting the choice to major in DSA may be the focus on entrepreneurial training as well as the senior-year capstone.

We enthusiastically support the efforts of the SNS, as we foresee there being various ways in which their efforts dovetail with the draft DSA major proposal that we have also shared, together advancing this important data science paradigm across campus.

Based upon their draft proposal, below we identify a few candidate areas to integrate efforts around data science education:

- Co-listing electives
- Co-listing certain complementary core requirements.
- Shared support for a data science education technician, for example to aid with maintaining JupyterHub cloud-based computing infrastructure to support scaling computational resources beyond what is available to students using their own laptops.

1.7.5 Relation to the UC Berkeley data science major

The UCB program is to a large degree the model around which our program was designed, whereby all students take common courses developing core data science skills, but specialize in one domain by way of select lower and upper-division course elective groups.

The proposed DSA major is also a collective effort that integrates existing faculty and curriculum to define broad application domains that largely reflect the four UC Merced departments involved: Cognitive & Information Sciences, Economics, Management of Complex Systems, and Political Science. As such, the application domains in the proposed DSA major map onto roughly seven of the 29 UC Berkeley specialization domains: Business and Industrial Analytics; Cognition; Economics; Environment, Resource Management, and Society; Geospatial Information and Technology; Organizations and the Economy; Science, Technology, and Society; Social Policy and Law. This leaves 22 domains undeveloped, providing ample room at UC Merced to develop data science programs in the School of Natural Sciences and School of Humanities, Social Sciences and Arts that address appropriate corresponding domains of faculty expertise that leverage existing curriculum.

For additional comparison of the UC Berkeley data science fundamentals courses with the corresponding courses in the DSA curriculum, see the Appendix Documents folder file "**DSA_Course-level Relation to the UC Berkeley data science major.docx**". In summary, regarding the mastery of technical computing and quantitative skills, the proposed DSA major progresses over a longer time frame than the UCB program, which may appeal to a different set of students.

2. Program Requirements

2.1 Overview

The major coursework consists of two types of courses: (i) DSA core and principles, which develop technical skills (e.g. coding in scripting languages, visualization of multi-level data, statistical inference); and (ii) application domains electives, where candidate problem domains are introduced and technical skills can be deployed in order to solve a specific problem.

In an effort to emulate the success of the culminating capstone incorporated by the School of Engineering, this program will also provide a year-long experience that launches students into the domain of real-world problem solving under time, team and other resource constraints. The senior capstone in its simplest form is a team-based research project supervised by the instructor, in which the target problem is inspired or derived from a real-world problem. (See Section 4.1.1 for detailed information on the administration of capstone.)

As the program matures, local (campus administrators, faculty) and regional (public, private businesses) partners and clients will have the opportunity to provide target problems that are appropriate for either a single team, or are sufficiently oriented such that the project could be passed along cohorts, which was a successful model used by the Engineering Service Learning program at UC Merced.

2.2 Lower and upper division course requirements

The majority of required DSA courses already exist. Only one new course (DSA 002) would be needed to start the program in AY 2024-25; and in total, only four courses (DSA 002; 101 or 102; 103; 120/121) would need to be introduced for the first cohort to graduate in AY 2028-29.

2.2.1 Course Requirements:

Major – 17 Courses, 64 Units

Table 2: Distribution of required courses

		Courses	Units
Lower Division			
	Major Preparation	4	16
	DSA Core	4	16
Upper Division			
	DSA Principles	4	16
	Application Domain Electives	3	12
	Capstone	2	4
	Total	**17**	**64**

In the lists provided below, new (to-be-proposed) courses are indicated with an asterisk (*). See *Appendix I* for course descriptions.

Lower Division Major Requirements [8 courses, 32 Units]

Required Major Preparation [4 courses]

All Data Science and Analytics majors are required to complete the following lower-division preparation courses.

Mathematics Requirement [8 units]

Complete the following course:

- MATH 011 or MATH 021: Calculus I (Units: 4)
- MATH 012 or MATH 022: Calculus II (Units: 4)

Foundation Requirement [8 units]

Complete the following courses:

- ECON 001: Introduction to Economics (Units 4) or COGS 001: Introductions to Cognitive Science (Units 4)
    - Or apply for waiver to accept 1st year "Social Science GE" if CSE Transfer student
- PHIL 002: Introduction to Ethics (Units 4) or PHIL 003: Contemporary Moral Problems (Units 4)

Data Science and Analytics Lower Division Core Requirement [4 courses, 16 units]

Complete the following courses:

- DSA 001: Foundations of DSA (to be cross-listed with existing course, MIST 060: Introductory Data Analytics) (Units: 4)
- * DSA 002: Introduction to Computing and Logic (Units: 4)
    - Transfer student can substitute with CSE 019 Intro to Computing (Units: 4) or CSE 022 Intro to Programming (Units: 4).
- ECON 010: Statistical Inference (Units: 4)
- MIST 050: Introduction to Entrepreneurship (Units: 4) or * MIST 070: Creativity and Innovation (Units: 4)

Upper Division Major Requirements [9 courses, 32 Units]

Upper Division Data Science and Analytics Principles Requirement [4 courses, 16 units]

Complete the following courses:

- ECON 110: Econometrics (Units: 4)
- * DSA 103/COGS XXX/PHIL 1XX: Data Ethics (Units: 4)
- Advanced DSA Methods A – choose one from:
    - MIST 130: Statistical Data Analysis and Optimization in R for Decision Support (Units: 4) or
    - MIST 134: Methods of Data and Network Science (Units: 4)
- Advanced DSA Methods B – choose one from:
    - * DSA 101: Machine Learning & Natural Language Processing (Units: 4) or
    - * DSA 102: Interactive Data Visualization (Units: 4) or
    - MIST 135 (existing course; to be cross-listed as DSA 104): Technical Communication and Visualization (Units: 4)

Data Science and Analytics Capstone Requirement [2 courses, 4 units]
Complete the following two courses:

- * DSA 120 Capstone 1 (Units: 2)
- * DSA 121 Capstone 2 (Units: 2)

Application Domain Elective [3 courses, 12 units]
Complete at least three additional courses listed below (at most one lower-division prerequisite listed along the courses below can be counted toward the three courses), which are grouped by application domains for convenience. Note that CSE 024: Advanced Programming is only allowed for UCM CSE transfer students.

The Environment and Sustainability

- ENVE/MIST 110: Hydrology and Climate (pre-req. includes: ENVE 020 or MATH 015)
- ESS/BIO 113: Sustainability in the Anthropocene
- MIST 116/ESS 132: Applied Climatology (3 units; pre-req.: ESS 110 or ENVE 110)
- MIST 118/ESS 1XX: Climate Change: Science & Solutions
- ENGR 141: Environmental Science and Policy
- MIST 132: Geographic Information Systems Analysis in Management
- PHIL 122: Bioethics
- MIST 120 Parks and Protected Areas
- ENVE 130 Meteorology and Air Pollution (prereq: ENVE 020 or ESS 020)
- ENVE 152 Remote Sensing of the Environment (prereq: MATH 021 or exam, and PHYS 008 or exam)

Policy and Decision Making

- ECON 120: Economics of the Environment and Public Policy
- ECON 153: Judgement and Decision Making
- MIST 131: Data Governance for Analytics Projects
- POLI 112: Public Policy: Analysis, Strategy and Impact
- POLI 120: Voting Behavior, Campaigns and Elections
- POLI 170: Theoretical Models of Politics
- POLI 175: Advanced Analysis of Political Data

Understanding and Modeling People

- MIST 138: Systematic Financial Trading & Analysis
- CSE 175/COGS125: Intro to Artificial Intelligence
- COGS 103: Intro to Neural Networks in Cognitive Science
- COGS 104: Complex Adaptive Systems
- COGS 105 Research Methods for Cognitive Science (pre-reqs. include PSY 010)
- COGS 122: Modeling Social Behavior
- COGS 128: Cognitive Engineering
- COGS 130: Cognitive Neuroscience
- COGS 170: Judgement and Decision Making

Other Elective Courses

- CSE 024: Advanced Programming (Can be counted only if CSE transfer)
- ENVE 155: Decision Analysis in Management (pre-reqs. include ECON 100)
- MIST 130: Statistical Data Analysis and Optimization in R for Decision Support
    - Allowed as elective only in the case that MIST 134 is also taken to satisfy the Advanced Methods Course requirement
- MIST 134: Methods of Data and Network Sciences
    - Allowed as elective only in the case that MIST 130 is also taken to satisfy the Advanced Methods Course requirement
- MIST 175: Information Systems for Management
- MIST 1XX (planned): User Experience
- MIST 190: Special Topics Course
- Any of the upper-division Advanced DSA courses beyond the single required course

2.3 Learning goals and outcomes

Upon completion of the program, students will have the skillset needed to develop a workflow for harnessing and making use of data in a range of real-world contexts and be prepared to engage in lifelong learning and continuing education by leveraging this skillset in either industry or academic future pathways.

Five PLOs for the proposed DSA Major:

1. **Methods.** Students will be able to identify data-oriented solutions for problems associated with evaluation, management, planning and strategic decision-making by integrating statistical, data science, analytical thinking and visualization methods.

2. **Communication**. Students will strategically use 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. **Application Design.** Faced with multi-dimensional problems that can span multiple boundaries of business, management, economics, sociology, psychology, cognitive science, environmental science and engineering, students will know how to apply design principles to develop an ethical and feasible data-driven strategy that strategically collects, manages, analyzes, communicates, and delivers information-based services.

4. **Team.** Students will leverage productivity software for project planning and team collaboration, effectively and efficiently integrating team members' diverse skillsets and knowledge in professional high-stakes scenarios with time and other resource constraints.

5. **Ethics.** Students will be able to identify ethical, legal, and social issues surrounding data collection and analysis; and to creatively develop and evaluate responses to these issues; and to implement appropriate responses.

2.4 Course relationship with learning outcomes

How the lower and upper division course requirements support the 5 PLOs is illustrated in **Table 3**. The second row lists the key competencies and skillsets for each PLO that are commonly listed in corresponding job announcements (see *Appendix II - IRDS*). The bottom row indicates the year that assessment of the given PLO will occur.

Table 3: Curriculum map relating required courses to PLOs and their relevant skills.

COURSE / PLO	PLO 1: METHODS	2: COMMUNICATION	3: APPLICAT-ION DESIGN	4: TEAM	5:ETHICS
SKILLS	(Tabular) Microsoft Excel and non-tabular (Python/SQL/Linux/Git) computing environments; Data Science; Machine Learning; Business Intelligence	Communication Skills; (Technical) Writing; Detail-oriented; Tableau; Data Visualization;	Problem Solving; Troubleshooting; Research; Creativity; Data Science; Machine Learning; Business Intelligence; Economics	Teamwork / Collaboration; Planning; Microsoft Productivity / Teams suite	Data Science
MATHEMATICS AND COMPUTER SCIENCE REQUIREMENT					
MATH 011/012: CALCULUS I/II	X				
FOUNDATION REQUIREMENT					
ECON 001: INTRODUCTION TO ECONOMICS			X		
COGS 001: INTRODUCTIONS TO COGNITIVE SCIENCE			X		
PHIL 002 OR PHIL 003					X
DATA SCIENCE AND ANALYTICS LOWER DIVISION CORE REQUIREMENT					
DSA 001/MIST 060: FOUNDATIONS OF DSA	X	X			
DSA 002: INTRO TO COMPUTING AND LOGIC	X		X		
ECON 010: STATISTICAL INFERENCE	X		X		
MIST 050: INTRO TO ENTREPRENEURSHIP		X	X	X	X
UPPER DIVISION DATA SCIENCE AND ANALYTICS PRINCIPLES REQUIREMENT					
ECON 110: ECONOMETRICS	X		X		
MIST 135: TECHNICAL COMMUNICATIONS AND VISUALIZATION		X	X	X	
MIST 130: STATISTICAL DATA ANALYSIS AND OPTIMIZATION IN R FOR DECISION SUPPORT	X		X	X	
MIST 134: METHODS OF DATA AND NETWORK SCIENCE	X	X			
COGS XXX/PHIL 1XX/DSA 103: DATA ETHICS			X	X	X
DSA 101: MACHINE LEARNING & NATURAL LANGUAGE PROCESSING	X		X		X
DSA 102: INTERACTIVE DATA VISUALIZATION	X	X			X
DSA 104/MIST 135: TECH. COMM. AND VISUALIZATION		X	X	X	
DATA SCIENCE AND ANALYTICS CAPSTONE REQUIREMENT					
DSA 120 - FALL	X	X	X	X	X
DSA 121 - SPRING	X	X	X	X	X
ASSESSMENT					
YEAR (COURSE)	1 (DSA 001), 3 (MIST 130), 4 (Capstone)	2 (DSA 001), 3 (MIST 135), 5 (Capstone)	2 (MIST 070), 5 (DSA 101), 4 (Capstone)	3 (MIST 050), 6 (Capstone), 4 (MIST 130)	4 (DSA 103) 5 (Capstone) 6 (MIST 050)

2.5 Outcome assessments

We plan to follow standard UCM protocol for administering program assessment associated with PLOs. Annually, direct and indirect evidence of student learning will be obtained and administered by relevant instructional faculty in conjunction with the Academic Planning and Assessment Support. Indirect evidence is obtained via surveys (such as the annual Graduating Senior Survey and biannual UC Undergraduate Experience Survey) and, as need arises, focus groups and interviews. Direct evidence is obtained from student work in core courses. The last row of **Table 3** indicates the year in which each PLO will be evaluated, along with the specific course. **Table 4** shows the PLO-specific rubrics that will be used to evaluate specific assignment or project outcomes according to four performance levels, thereby identifying areas for improvement and ways in which different courses addressing the same PLO can support each other. These annual assessments will be incorporated into the periodic (6-year interval) review of the program.

Programmatic continuity will be assessed by way of tracking the number of students declaring the major upon arrival to UC Merced, and tracking the retention rate by cohort and program year. Once the program is in year 3, similar tracking of transfer students into the program from other majors and from community colleges will provide analog measures of retention.

Periodic program review assessment based upon the afore-mentioned annual PLO assessments and exit reviews will evaluate: (a) program/faculty quality and support; (b) student learning outcomes and achievements; (c) capstone outcomes; (d) employment and post-graduate outcomes, either communicated by students or by tracking alumni LinkedIn profile updates.

Table 4. PLO Assessment Rubric

PLO/Level	Expert	Advanced	Intermediate	Introductory
1. Methods	Work demonstrates a specific combination of data-oriented methods to optimally address a specific problem. Able to integrate a specific data source, statistical method, experimental design, and visualization method to provide an appropriate solution. Description of the implementation plan is clear and understandable by others, and is accompanied by working code demonstrating the implementation.	Work demonstrates how to combine data-oriented methods to address a specific problem. Able to integrate a relevant data source, statistical method, experimental design, and visualization method to make progress towards a solution. Description of the implementation plan is clear and understandable by others.	Work demonstrates basic understanding of how to arbitrarily combine data-oriented methods to address a specific problem. Unable to integrate one of the following - data source, statistical method, experimental design, or visualization method – and so the solution pathway is incomplete. Description of the implementation plan is not clear, unfeasible or misguided.	Unclear how to use any combination of data-oriented methods to address a specific problem. Unable to identify a relevant data source, statistical method, experimental design, or visualization method. Description of the implementation plan is not clear, unfeasible, or altogether lacking. Attempt to use a preliminary method results in getting stuck.

2. Communication	Writing effectively communicates with data, clearly documents data source, method & assumptions, and technical terminology; is engaging and inspires further inquiry; and is concise and well-aligned with the expectations and technical background of the target audience. Supporting statistical analysis and/or data visualizations are designed to capture attention and imagination, while also fully supporting the thesis; and are tailored to the particular audience. Interactive data visualizations provide the audience the opportunity to interact and explore the supporting analysis remotely. Oral presentation is clear, engaging and inspirational. Presentation is organized, includes eye-catching data visualization, and satisfied allotted time constraints. Written code is well organized and it is possible for others to understand and execute it themselves. Code is of sufficient quality to be shared on a public repository.	Writing effectively communicates with data, clearly documents data source, method & assumptions, and technical terminology; but is lacking an overall inspirational objective, or is not concise. Supporting statistical analysis and/or data visualizations are presented; are sufficient to fully support the thesis; are tailored to the particular audience. Oral presentation is clear and professional grade. Presentation is organized, logical in its flow, includes original data visualization that is polished, and satisfies allotted time constraints. Written code is well organized and it is possible for others to understand and execute it themselves.	Writing suffers in just a single area, is not entirely clear in terms of definitions and assumptions, and does not effectively inform, engage or inspire in a clear and concise manner. Disjoint integration of data. Supporting statistical analysis and/or data visualizations are presented, but are flawed in their assumptions; are otherwise misleading; are insufficient to fully support the thesis; or are not fully designed for the particular audience. Oral presentation is understandable but it is difficult to understand the objective of the communication. Presentation is unorganized, includes original data visualization that is unpolished, or goes over the allotted time. Written code is poorly organized and hard for others to understand.	Writing is unclear and does not effectively inform, engage or inspire in a clear and concise manner. Supporting statistical analysis and/or data visualizations are presented, but do not inform in any particularly strategic way towards the communication thesis. Oral presentation is ineffective and it is difficult to understand the objective of the communication. Presentation is unorganized, lacks original data visualization, or goes over the allotted time. Written code is not organized, does not employ comments, and/or contains critical syntax errors.

3. Application Design	The problem-solving approach identifies feasible strategy for developing a marketable application. Several feasible alternative approaches are also identified and ranked. The proposed solution optimally balances across ethical, legal and social factors that constrain the problem-solving space. The proposed solution has boundary-spanning applications and can be scaled to reach across multiple application domains.	The problem-solving approach identifies a feasible strategy for achieving a well-defined objective. Several alternative approaches are also identified and considered. The proposed solution sufficiently balances across ethical, legal and social factors that constrain the problem-solving space. The proposed solution has wide application scope.	The problem-solving approach identifies both an objective and strategy for achieving the objective, but lacks consideration of feasibility. The proposed solution casually considers ethical, legal and social factors that constrain the problem-solving space. The proposed solution has limited application scope.	The problem-solving approach is piecemeal, lacking a clear objective, lacking a clear strategy, and lacking alignment between the strategy and the objective. The proposed solution lacks consideration for ethical, legal and social factors that constrain the problem-solving space. The proposed solution has no real-world utility.

4. Team	Tasks are equitably delegated and leverage the strengths of the team members. All team members attend meetings and thoroughly complete assigned work on schedule. Delegated tasks are tracked and supported through group accountability. All team members maintain responsibility for their parts and awareness of all parts of project. Team members build on the work of others to advance the project. Team members identify and acknowledge conflict and address it appropriately. All team members make strong contributions to data analysis and coding.	Tasks are equitably delegated and leverage the strengths of the team members. All team members attend meetings and thoroughly complete assigned work on schedule. Team members identify relevant, alternative solutions or ideas that build on the work of others to advance the project. Team members identify and acknowledge conflict and address it appropriately. All team members contribute even incrementally to data analysis and coding.	Tasks are delegated disproportionately across team members. Team members attend meetings and completed assignments are mostly submitted on time and mostly thorough. One or two team members offer new ideas and seem to drive the group while the others contribute by expressing support for those ideas. Conflict is acknowledged but not addressed appropriately. A limited number of students are responsible for data analysis and coding.	Tasks are unevenly distributed across team members. Assigned work is turned in on time, but is not always complete. A small number of team members dominate the group and are responsible for data analysis and coding, while the others passively follow along. Conflict is not acknowledged.
5. Ethics	Identifies, critically evaluates, and accounts for multiple specific ethical, legal and/or social issues related to data collection, methodology and/or application. The ramifications of a particular methodology or application and the plausibility for unintended negative consequences, both in the short-term and long-term, are elaborated.	Solid understanding of ethics and its importance, and able to identify multiple specific ethical, legal and/or social issues related to data collection, methodology and/or application. The ramifications of a particular methodology or application and the plausibility for unintended negative consequences are elaborated.	Solid understanding of ethics and its importance, and able to identify at least one specific ethical, legal or social issue related to data collection, methodology or application. The ramifications of a particular methodology or application and the plausibility for unintended negative consequences are clearly considered.	Able to understand importance of ethics, but unable to identify a specific ethical, legal or social issue related to data collection, methodology or application. A thorough discussion of the ramifications of a particular methodology or application and the plausibility for unintended negative consequences is vague, or missing.

2.6 Minimum and maximum credits

Across all courses: Range: 96-108 (Minimum assumes a student testing out of MATH 011, 012 and ECON 001 based on sufficient AP test score).

Focusing within the subset of required program courses: Range: 52-64 (Minimum assumes a student testing out of MATH 011, 012 and ECON 001 based on sufficient AP test score; Maximum corresponds to 17 required courses).

2.7 Sample Program

The lower-division course requirements are similar to other STEM majors in that students must pass fundamental math courses (Calculus I and II) to satisfy this pre-requisite for upper-division STEM courses. Students who fail to pass these courses will have difficulty in completing the major in 4 years, but this would also be the case in many STEM majors.

Once students complete the lower-division courses that are common pre-requisites for other upper-division courses, there is flexibility built into the program by way of 3 PPP electives, which provide ample room for student modification and course correction.

Table 5 shows the feasibility of accomplishing this major in 4 years if the student declares the major in Y1. **Table 6** shows the straightforward integration of a CSE major who transfers after completing the standard first-year CSE curriculum. Because many first-year courses map exactly onto the DSA first year, such a student could be fully integrated with their cohort staring in Y2. For a CSE major who transfers after the second year, the course timeline would look very similar to what is shown in **Table 7** for community college transfer, however without DSA 001 (satisfied by CSE 022 or MATH 032), DSA 002 (satisfied by CSE 015) and one PPP Elective (satisfied by CSE 024).

Table 5: Example 4-year Student Course Calendar:

The 17 required DSA courses are colored in **blue** and **green**.

Year 1: Fall - 16 units	Year 1: Spring - 16 units	Year 2: Fall - 16 units	Year 2: Spring - 16 units
Math 011: Calculus I	Math 012: Calculus II	MIST 050: Intro to Entrepreneurship or PPP Elective	GE/Elective
*DSA 002: Intro to Computing and Logic	ECON 001 / COGS 001	GE/Elective	ECON 010: Statistical Inference
DSA 001: Foundations of DSA	WRI 010: College Reading and Composition	GE/Elective	GE/Elective
SPRK 001: Spark Seminar	PHIL 002 / PHIL 003	GE/Elective	GE/Elective

Year 3: Fall - 16 units	Year 3: Spring - 12 units	Year 4: Fall - 14 units	Year 4: Spring - 14 units
ECON 110: Econometrics	*COGS/PHIL/DSA 103	Advanced DSA Meth. B: *DSA 101 or *DSA 102 or MIST 135/DSA 104	GE/Elective
Advanced DSA Meth. A : MIST 130 or MIST 134	PPP Elective	*DSA 120: Capstone I	*DSA 121: Capstone II
GE/Elective	GE/Elective	PPP Elective	PPP Elective
GE/Elective		GE/Elective	GE/Elective

Major Preparation Course (4 total)
DSA Course Requirement (13 total)
GE or Elective Course
* New Course

Table 6: Example 4-year Student Course Calendar

UCM CSE Transfer after completing first year:
CSE 022 counts for DSA 001; CSE 015 counts for DSA 002; CSE 024 counts for a PPP elective.
The 17 required DSA courses are colored in **blue** and **green**.

Year 1: Fall - 16 units	Year 1: Spring - 16 units	Year 2: Fall - 16 units	Year 2: Spring - 16 units
Math 021: Calculus I	Math 022: Calculus II	MIST 050: Intro to Entrepreneurship or PPP Elective	PHIL 002 / PHIL 003
CSE 022: Intro to Programming (Satisfies DSA 001)	Social Science GE	GE/Elective	ECON 010: Statistical Inference
WRI 010: College Reading and Composition	CSE 024: Advanced Programming (Satisfies PPP Elective)	ECON 001 / COGS 001	GE/Elective
SPRK 001: Spark Seminar	CSE 015: Discrete Math (Satisfies DSA 002)	GE/Elective	GE/Elective

Year 3: Fall - 16 units	Year 3: Spring - 12 units	Year 4: Fall - 14 units	Year 4: Spring - 14 units
ECON 110: Econometrics	*COGS/PHIL/DSA 103	Advanced DSA Meth. B: *DSA 101 or *DSA 102 or MIST 135/DSA 104	GE/Elective
Advanced DSA Meth. A : MIST 130 or MIST 134	PPP Elective	*DSA 120: Capstone I	*DSA 121: Capstone II
GE/Elective	GE/Elective	PPP Elective	GE/Elective
GE/Elective		GE/Elective	GE/Elective

Major Preparation Course (4 total)
DSA Course Requirement (13 total)
GE or Elective Course

* New Course

Table 7: Example Transfer Student Course Calendar

MCC Transfer after completing two years; the course calendar for 2nd year UCM CSE transfer would look similar, but would already have satisfied DSA 001, 002 and one PPP Elective. The 17 required DSA courses are colored in **blue** and **green**.

Example: Merced College Courses			
Fall - 14 units	Spring - 16 units	Fall - 16 units	Spring - 16 units
ENGL 1A (4 units)	ENGL 13 (3 units)	CPSC 07 (3 units) [DSA 002 substitute]	Elective/Ex: PHIL 05 (3 units)
MATH 04A (4 units)	ECON 2 (3 units)	MATH 10 (3 units)	Elective/Ex: ANTH 01 (4 units)
CPSC 06 (3 units)	MATH 04B (4 units)	PHIL 02 (3 units)	Elective/Ex: POSC 01 (3 units)
ECON 1 (3 units)	Elective/Ex: HIST 17A (3 units)	Elective/Ex: GEOL 01 (4 units)	Elective/Ex: PSY 01A (3 units)
	Elective/Ex: SOC 01 (3 units)	Elective/Ex: HIST 17B (3 units)	Elective/Ex: ARTD 40A (3 units)

Courses Taken at UC Merced			
Year 1: Fall - 16 units	Year 1: Spring - 16 units	Year 2: Fall - 14 units	Year 2: Spring - 14 units
DSA 001: Foundations of DSA	ECON 010: Statistical Inference	PPP Elective	GE/Elective
MIST 050: Intro to Entrepreneurship	PPP Elective	Advanced DSA Meth. B: *DSA 101 or *DSA 102 or MIST 135/DSA 104	PPP Elective
ECON 110: Econometrics	*COGS/PHIL/DSA 103	*DSA 120: Capstone I	*DSA 121: Capstone II
Advanced DSA Meth. A : MIST 130 or MIST 134	GE/Elective	GE/Elective	GE/Elective

2.8 Accessibility for transfer students

Table 7 shows the timeline for a hypothetical incoming third-year transfer student from Merced Community College accomplishing this major in two years after transferring. A substantial number of first and second-year courses can be completed via transfer credits from Merced College – i.e., MATH 011/012 (substituted by MATH 04A, 04B and 10); DSA 002 (CPSC 06 and 07), ECON 001/ 010 (Econ 1 and 2); PHIL 002/003 (PHIL 02). We will adopt the CC transfer protocol used by CIS for COGS 001, which does not have a natural counterpart.

One challenge for transfer students would be to satisfy the requirements of DSA 001, which is unique to the campus. One other lower-division courses that would be difficult to satisfy with transfer credits is MIST 050 (Entrepreneurship). Nevertheless, a dedicated student could take DSA 001 and MIST 050 (or MIST 070) in Year 3 or 4 as there are no pre-requisites.

Transfer students are expected to satisfy the minimal UC Merced requirements for the School of Engineering (see https://admissions.ucmerced.edu/transfer/major-preparation-schools - soe) until the Gallo School of Management has stipulated corresponding transfer requirements (e.g. min. GPA 2.4 for residents, and 2.8 for nonresidents, with a grade of "C" or better, and a minimum total of 60 UC transferable semester units the term prior to enrollment).

2.9 Draft Catalogue Description

Data Science and Analytics Major terminating in B.A.

The Data Science and Analytics (DSA) major addresses the increasing demand for adept data analysts by combining coursework in applied data science with coursework in economics and cognitive, political and management science – areas where real-world problem solving increasingly relies on data-driven analysis to inform decision-making.

The reason for the increasing workforce demand is the ubiquity of data, which are generated everywhere, and increasingly envelop our human experience by mediating our interactions with the natural, social and built world. Yet despite their prevalence, data are not necessarily generated in a way that are immediately usable or valuable. Instead, they require refinement, reformatting and critical analysis before they can be converted into something informative. In order to extract value that an individual or organization might use in order to make a decision – i.e., raise or lower a tax, hire or dismiss within a department, or invest or divest in a particular sector – one needs to understand how to appropriately and ethically apply methods of data science for harnessing and classifying the data, along with methods of data analytics for extracting actionable insights and effectively communicating the results.

The DSA major trains students across the entire data analytics pipeline consisting of series of refinements including, but not limited to: identifying and integrating appropriate data; understanding the strengths and weaknesses of the data and its sources; extracting valuable insights by way of advanced visualization and inferential methods; and strategically communicating recommendations based upon that analysis that maximally inspire some person or group to take action. A defining feature of this major is the senior-year capstone, where students will work in teams to apply advanced methods of interactive data visualization and machine-learning to address real-world problems. As such, this major provides students a unique opportunity to develop a multi-dimensional skillset – one that integrates core quantitative and computational capabilities with a broad understanding of societal context, ethical best practices, and real-world applications.

3. Accreditation

This program will not seek accreditation. As data science education is a new paradigm, to the best of our knowledge, there is no accreditation for this type of educational program.

4. Resources Needed and Plan for Providing Them

4.1 Faculty members

Table 8 shows the teaching plan and faculty involved in the first 5 years of the program. This program draws on existing faculty, and calls for no new faculty lines to launch or support the program long term. The following faculty have either taught courses in the past that are major requirements or electives, or have been consulted regarding the preparation of new courses.

CIS department: Hanna Gunn, Dan Hicks, Christopher Kello, David Noelle, Lace Padilla, Paul Smaldino, Jeffrey Yoshimi.

ECON department: Andrew Johnston.

MCS department: John Abatzoglou, Spencer Castro, Jeffrey Jenkins, Tea Lempiala, Russ McBride, Alex Petersen, LeRoy Westerling, Lisa Yeo.

Table 8: Five-year teaching plan

* indicates an UG service course that is already taught, independent of the proposed DSA major

** GS indicates a qualified graduate student who has passed to candidacy.

COURSE	YEAR 1	Y2	Y3	Y4	Y5
MATHEMATICS AND COMPUTER SCIENCE REQUIREMENT					
MATH 011/012: CALCULUS I/II	*	*	*	*	*
FOUNDATION REQUIREMENT					
ECON 001: INTRODUCTION TO ECONOMICS	*	*	*	*	*
COGS 001: INTRODUCTIONS TO COGNITIVE SCIENCE	*	*	*	*	*
PHIL 002 INTRO ETHICS OR PHIL 003	*	*	*	*	*
DATA SCIENCE AND ANALYTICS LOWER DIVISION CORE REQUIREMENT					
DSA 001/MIST 060: FOUNDATIONS OF DSA	Petersen / MCS / CIS	Petersen / MCS / CIS	Petersen / GS	MCS / CIS / GS	MCS / CIS / GS
DSA 002: INTRODUCTION TO COMPUTING AND LOGIC	McBride/Yeo	McBride/Yeo	McBride /Yeo/GS	McBride /Yeo/GS	McBride /Yeo/GS
MIST 050: INTRO TO ENTREPRENEURSHIP	McBride / Lempiala	McBride / Lempiala	McBride / GS	McBride / GS	McBride / GS
ECON 010: STATISTICAL INFERENCE	*	*	*	*	*
MIST 070: CREATIVITY AND INNOVATION		Lempiala / McBride	Lempiala / McBride	Lempiala / GS	Lempiala / GS
UPPER DIVISION DATA SCIENCE AND ANALYTICS PRINCIPLES REQUIREMENT					
ECON 110: ECONOMETRICS		* Johnston / Econ	* Johnston / Econ	* Johnston / Econ	* Johnston / Econ
MIST 130: STATISTICAL DATA ANALYSIS AND OPTIMIZATION IN R FOR DECISION SUPPORT		Castro / Westerling	Westerling / Castro	Castro / Westerling	Westerling / Castro
MIST 134: METHODS OF DATA AND NETWORK SCIENCE			Petersen / CIS		Petersen / CIS
COGS XXX/PHIL 1XX/DSA 103: DATA ETHICS		Gunn / Hicks	Gunn / Hicks	Gunn / Hicks	Gunn / Hicks
DSA 101: MACHINE LEARNING & NATURAL LANGUAGE PROCESSING				Petersen / Econ / POLISCI / CIS	Petersen / Econ / POLISCI / CIS
DSA 102: INTERACTIVE DATA VISUALIZATION AND MIST 135/DSA 104: TECH COMM. AND VISUALIZATION			Padilla /Petersen		Padilla / Petersen
CAPSTONE					
DSA 120 - FALL				Bhappu/Maglio	Bhappu/Maglio
DSA 121 - SPRING				Maglio/ Bhappu	Maglio/ Bhappu
RELEVANT ELECTIVES TAUGHT ON A REGULAR BASIS					

MIST 132: GEOGRAPHIC INFORMATION SYSTEMS ANALYSIS IN MANAGEMENT		Jenkins / Kolden	Jenkins / Kolden	Jenkins / Kolden	Jenkins / Kolden
MIST 175: INFORMATION SYSTEMS FOR MANAGEMENT			Yeo / Maglio		Yeo / Maglio
COGS 122: MODELING SOCIAL BEHAVIOR			Smaldino		Smaldino
CSE 175/COGS125: INTRO TO ARTIFICIAL INTELLIGENCE		Noelle		Noelle	
COGS 103: INTRODUCTION TO NEURAL NETWORKS IN COGNITIVE SCIENCE			Yoshimi		Yoshimi
MIST 116: APPLIED CLIMATOLOGY		Abatzoglou		Abatzoglou	

4.1.1 Administration of Senior Capstone

We do not expect this course to be anywhere near as resource-intensive to operate as the Engineering Capstone. All projects will be primarily data oriented, and hence requires few physical resources. Because no travel is required, this course is just two credits each semester. In general, because of the relatively low cost of data science and analytics resources (i.e., data and computing), we expect the financial resources required to launch and sustain this capstone will be significantly less than the resourced required of the Engineering School's Innovate2Grow, which involves hardware purchasing, lab space for mock-ups, and travel.

The course will be managed by a single instructor who will play an integral role matching students to problems, providing guidance and helping teams avoid pitfalls, facilitating meetings with professionals as needed, managing local web server access for testing and deployment of web-based tools, and organizing the final showcase in the Spring semester. The instructor will be identified by the Program Chair among the program faculty based on individual teaching load for the year .

4.2 Specialized staff/needed FTE

None.

4.3 Instructional space needs

None.

4.4 Instructional computing resources

This program will not require specialized hardware or software computing resources.

In the absence of a campus-wide laptop policy, the program will adopt the School of Engineering Policy on Laptops (effective Aug. 2019), such that students will be expected to have their own personal laptop suitable for computational analysis (e.g. minimum 16 GB RAM which has become the standard for Apple laptops). Otherwise the program will adopt the campus-wide laptop requirement that is under development.

Any classroom will suffice, as no specialized computing facilities will be needed for laboratory sections. This student-oriented approach to computing resources will accelerate learning and retention of data and applications management. It will also relieve the campus of limited instructional lab availability. We believe this approach will also accelerate practical skills

relating to virtual and cloud computing, and would complement efforts on campus to move in this direction to address the lack of instructional computing labs.

4.5 Library resources needed

This program will not require special library resources beyond what is already provided to undergraduate students in other STEM majors.

Regarding new Library/Collections needs, this program does not anticipate any new journals or data collections, as it will leverage open data sources and existing open textbook collections. Included in the Appendix is a letter of support on behalf of the UCM Library.

4.6 Resource needs for field studies or off-campus activities

None.

4.7 Specialized facilities

None.

5. Program Governance

The proposed DSA program will be offered jointly by four departments – CIS, EBM, MCS, and POLISCI – which currently span two schools, and are going through the process of joining together to form the future E&J Gallo School of Management.

MCS Governance: Prior to the Gallo School's approval, the DSA program will be administered by the MCS department. This choice has been agreed upon by the four departments, as it simplifies the program's administrative burden (due to the departments spanning two schools). This choice is reasonable given that MCS faculty offer a significant number of courses associated with the DSA program, many of which have been offered for some time through the Management Analytics and Decision-Making (MAD) minor.

Consequently, the proposed DSA major program was submitted by the MCS faculty for review by the School of Engineering Executive Committee before it was passed on for campus-level review. Based upon this review and subsequent internal discussion, the principal need identified by the School of Engineering is for sufficient student advising staff to support the DSA major.

During the interim period when the program is governed by MCS, there will be a program chair nominated by the department chairs of the four departments, in consultation with their faculties, and appointed by the Dean of the School of Engineering. This program chair will be responsible for all aspects of governance of the program in consultation with the four department chairs and graduate group chairs. For example, teaching assistants for DSA courses will be drawn from CIS/EBM/MCS/POLISCI-administered graduate programs, and possibly from other STEM graduate programs as needed. Priority for TA assignments will be determined by the home department of the faculty teaching specific courses in the major. For example, if an ECON faculty member is the instructor of record for DSA 001, the ECON department chair will identify an appropriately prepared TA among its graduate students; of course, if there is a shortage of ECON graduate students, then other Gallo School departments will be solicited to provide a TA.

Joint Governance: Once the Gallo School is formed, it will assume the student advising staff responsibilities, and the responsible dean will shift to the Dean of the Gallo School. In addition, the DSA program will be governed jointly by way of a representative curriculum committee that includes faculty from each of the four Gallo School departments, CIS, EBM, MCS, and POLISCI. DSA governance will be led by program chair that rotates every three years around the participating faculty departments. This chair will lead committees (comprised of representatives from each participating department) to address program development, review, and course-instructor assignments and planning. See Appendix IV for draft bylaws for DSA program governance (which are adapted from the bylaws of the Biological Sciences Undergraduate Program, the only other cross-department major on campus).

If the Gallo School is not formed within the first three years of operation of the DSA program, a joint governance model will be established formally (and ultimately approved by majority vote of the faculty in each of the four departments), following roughly the draft bylaws in Appendix IV, but allowing for cross-school operation.

6. Potential for non-majors to participate

All courses except for DSA 120/121 (Capstone) will be open to non-majors satisfying the indicated prerequisites.

7. Timetable for implementation

7.1 Timeline

First cohort in Fall 2024 – Freshman only. Only one new course, DSA 002 Intro to Computing and Logic, needed to launch the program for the first freshman cohort.

Sophomore-level UCM transfers (respectively, junior-level community college transfers) will be considered in Fall 2025 (resp., Fall 2026).

Appendix I – Course Descriptions

If no prerequisites are listed then there are none.

Lower Division – 8 Courses, 32 Units

Major Preparation Course – 4 Courses, 16 Units

MATH 011*: Calculus I* (4 Units) - Introduction to differential and integral calculus of functions of one variable, including exponential, logarithmic and trigonometric functions, emphasizing conceptual understanding and applying mathematical concepts to real-world problems (approximation, optimization). Pre-req: MATH 005 or equivalent exam.

MATH 012*: Calculus II* (4 Units) - Continuation of MATH 011. Introduction to integral calculus of functions of one variable and differential equations, emphasizing conceptual understanding and applying mathematical concepts to real-world problem. Pre-rq: MATH 011 or 021 or eq. ex.

ECON 001 or COGS 001

ECON 001*: Introduction to Economics* (4 Units) - Introduction to economics principles and methods, including microeconomics (operation of the economy at the individual and firm level) and macroeconomics (nature and functions of the national economy in a global context).

COGS 001*: Introduction to Cognitive Science* (4 Units) - An introduction to the interdisciplinary field of cognitive science. Basic issues related to cognition, including perception, memory, language, learning, problem solving, spatial cognition, attention, mental imagery, consciousness, brain damage, development, and artificial intelligence, are considered from the perspectives of psychology, philosophy, computer science, and neuroscience.

PHIL 002 or PHIL 003

PHIL 002*: Introduction to Ethics* (4 Units) - Consideration of basic questions and themes in moral philosophy through the study of historical and contemporary philosophical texts. Questions may include: Are moral rules relative? What is the best sort of human life? What sorts of acts are right? Should wealth be shared with those less fortunate?

PHIL 003: *Contemporary Moral Problems* (4 Units) - Consideration of topics in applied ethics, which may include euthanasia, abortion, economic justice, world hunger, the treatment of animals, and punishment. This course will also cover some more general approaches to ethical thinking, such as theories based on rights and entitlements.

DSA Lower Division Core Course – 4 Courses, 16 Units

ECON 010*: Statistical Inference* (4 Units) - Introduction to the application of social scientific methods to the study of economics, politics, and management. Covers research design, random sampling, descriptive and inferential statistics, hypothesis testing, and the linear regression model with an emphasis on applications. Pre-requisites: MATH 005 or MATH 011 or MATH 021 or equivalent exam.

***MIST 050 or MIST 070**:*

> ***MIST 050:** Introduction to Entrepreneurship* (4 Units) - Provides hands-on training in the practice of entrepreneurship where students learn to make actual money in challenging exercises in "the real world" (or measurably improve a social good if students choose a social entrepreneurship project).
>
> **** MIST 070:** Creativity and Innovation* (4 Units) – Introduces principles of creativity, innovation theory and methods of design thinking, in particular in team-oriented settings, for identifying strategies to address problems characterized by coupled social, human behavioral, technological, and environmental constraints.

**** DSA 001**: Foundations of Data Science & Analytics* (4 Units) - Introduces the data-for-decision-making paradigm through the lens of the Wolfram computing suite – comprised of WolframAlpha, the Mathematica language, and computable notebooks. Lectures introduce methods and best practices of data collection, integration and visualization, together reinforced via team exercises drawing on real-world domains. To facilitate explorative autodidactic learning, students will learn how to self-instruct new coding skills by using Mathematica Document Center 'help pages' in conjunction with myriad open coding demonstrations. Course provides a computational-oriented introduction to essential concepts of statistics, probability.

** This course will be cross-listed with existing course MIST 060 "Introductory Data Analytics", which is part of the existing MAD minor, and satisfies the GE Quantitative Reasoning badge and Approaches to Knowledge: Engineering Science. CSE transfer students transferring into the program will be able to satisfy this course requirement based upon satisfactory passing grade for CSE 022 or MATH 032.

**** DSA 002**: Intro to Computing and Logic* (4 Units) - Presents the basics of modern computing hardware, software and coding and fundamentals of discrete mathematics and logic theory for students with no prior experience. Students will learn principles of file system management, data file input/conversion/output, elementary coding with variables and multi-dimensional data structures, conditional logic statements and automated loops, debugging logic, open-source plotting packages, and package management using open-source scripting software installed on their own laptop.

> ** CSE transfer students transferring into the program will be able to satisfy this course requirement based upon satisfactory grade for CSE 015.

Upper Division - 9 Courses, 32 Units

DSA Principles - 4 Courses, 16 Units

***ECON 110**: Econometrics* (4 Units) – Introduction of problems of observation, estimation and hypotheses testing in economics through the study of the theory and application of linear regression models, critical evaluation of selected examples of empirical research and exercises in applied economics. Pre-Reqs: (ECON 010 or eq. ex.) and (MATH 011 or MATH 021 or eq. ex.).

** **COGS 1XX/PHIL 1XX/DSA 103**: Data Ethics (4 Units)* – An introduction to the ethical issues in the development and deployment of new technologies including big data, artificial intelligence, new media. Students learn about ethical theory, how to problem solve in ethics, and how to navigate the philosophical dilemmas that underlie ethical problems. Students will develop skills in applied ethical analysis and holistic thinking about complex systems, while developing a deeper understanding of social justice, autonomy, and responsibility. Pre-reqs: PHIL 002 or PHIL 003 or equivalent exam.

Advanced Method Courses:

Group A: Choose one from MIST 130 or MIST 134

***MIST 130**: Statistical Data Analysis and Optimization in R for Decision Support* (4 Units) – Introduces statistical analysis and optimization for decision support using the R programming environment. Analyze and visualize data and model relationships using graphing techniques, correlation and regression analysis, scenario development and analysis, sensitivity analysis, simulations, and optimization. Emphasizes applied work using real data from resource management and service management examples to support decision making. Case driven and team-focused, emphasizes best practices and professional ethics for analyzing, modeling and communicating empirical data and model results. Pre-Reqs: (MATH 032 or MATH 018 or BIO 018 or eq. ex.) and (MATH 050 or ME 021 or BIOE 021 or CSE 020 or eq. ex.). The Pre-requisites for this course will be reduced to resemble MIST 134.

***MIST 134**: Methods of Data & Network Science* (4 Units) - Explores methods to efficiently manage and analyze large complex datasets. The computer lab will introduce the Python programming language with a focus on extracting data from websites, exploring, and inquiry-oriented analyses and visualizations. Introduction to network science theory useful in understanding and managing complex socio-technological systems. Pre-Reqs: (MATH 012 or eq. ex.) and (DSA 002 or CSE 019 or ME 021 or BIO 018 or MATH 015 or BIOE 021 or eq. ex.).

Group B: Choose one from DSA 101 or DSA 102 or MIST 135/DSA 104

**** DSA-101**: Machine Learning & Natural Language Processing* (4 Units) – Lectures will introduces fundamental theoretical principles and applications of machine learning for the classification of text and images. The computer lab will develop intuition and best-practices of training set construction (labeling) and validation (sampling, robustness-checking), experimental design and inference. Case studies involving optical character recognition and semantic classification that draw on large-scale corpora will introduce students to Natural Language Processing principles relevant to the analysis of social/news media and policy documents. Pre-Reqs: (ECON 010 or BIO 018 or PSY 010 or by instructor approval) and (DSA 002 or CSE 019) and (MIST 130 or by instructor approval)

**** DSA-102:** Interactive Data Visualization* (4 Units) –Analyzing and presenting data visually has become one of the most essential skills for students who want to work in data science and related fields. Information Visualization teaches students how to design

compelling interactive visualizations of complex data to understand, discover, and present the information. The course combines theoretical knowledge and practical work to develop a well-rounded set of skills to ideate, design, implement, and evaluate sophisticated data visualization projects. The theoretical component aims to provide a mental model to think about the visualization design space in a principled manner. This includes the theory of visual encoding, human perception, and visualization techniques. The practical component aims to teach the skills needed to develop compelling interactive data visualizations for analysis and presentation. Pre-reqs: TBD.

MIST 135 (existing) / *DSA 104 (to be cross-listed): *Technical Communications and Visualization* (4 Units) - Theory of effective communication in quantitative contexts drawing on various presentation modes – written, oral, graphical, powerpoint, poster. Seminars on theory and best-practices combined with individual and team project presentations. Opportunity to improve inquiry-oriented communication skills, to give and receive constructive feedback, and to learn graphical methods for developing striking visualizations.

Application-domain (PPP) Electives – 3 Courses, 12 Units

The DSA curriculum supports gaining broad knowledge of data-oriented application domains by way of a flexible set of electives – requiring a minimum of 3 courses from a list of qualified application-domain courses, with at most one being lower-division. Students seeking to include courses not included in this list will be able to submit a written request to be considered for approval by the program's executive committee on a case-by-case basis. Only CSE transfer students can use CSE 024 to satisfy one PPP elective.

This framework facilitates students specializing in a particular application domain if they plan their electives carefully and early from Y1 onward. In some cases, students will need to identify the pre-requisite courses needed to reach a final destination course, e.g. taking MATH 021 and 022 instead of MATH 011 and 012. While we do not expect any or many students to identify specialized tracks, given the cross-disciplinary composition of the DSA faculty and curriculum, it is worth highlighting the potential for such ambitious schedules that are reminiscent of self-designed majors found at other universities.

Capstone – 2 Courses, 4 Units

**** DSA 120:*** *Capstone I* (2 Units) – Integrative team-based 2-semester Capstone project to address a real-world problem defined by an academic or industry advisor.

**** DSA 121:*** *Capstone II* (2 Units) – Continuation of DSA 120. Satisfies GE Culminating Experience badge. Pre-Reqs: DSA 120

Appendix II – IRDS

Job Market Assessment

For more elaboration, see Appendix Folder document: Data Science Job Types v2 12.17.2021.docx

Data Science

A program validation analysis using Burning Glass found that in the last 12 months in California 207,941 jobs were posted for graduates with bachelor's degree in Data Science. This represented 11% of jobs posted for graduates with bachelor's degree in the same period. The Burning Glass assessment estimated about 353,000 Data Science jobs in California in 2015, increasing to 384,000 Data Science jobs in 2020 and predicted to grow to 466,000 Data Science jobs in 2030, which is high relative to growth to growth in other fields. The Burning Glass assessment also found that median California salary for graduates with a bachelor's degree in Data Science would be $104,000.

Data Analytics

Applying the same analysis used for Data Science, Burning Glass found in the last 12 months in California 39,469 jobs were posted for graduates with bachelor's degree in Data Analytics. This represented 2% of jobs posted for graduates with bachelor's degree in the same period. The Burning Glass assessment estimated about 69,843 Data Analytics jobs in California in 2015, decreasing somewhat to 67,356 Data Analytics jobs in 2020 but predicted to grow to 76,135 Data Analytics jobs in 2030, which is high relative to growth in other fields. These data are echoed by the Bureau of Labor Statistics, which projects the employment of operations analysts to grow 25% between 2020 and 2030, a rate that is much faster than average. The Burning Glass assessment found that median California salary for graduates with a bachelor's degree in Data Analytics would be $85,000.

Data Science and Data Analytics

Applying the same analysis for both programs together, Burning Glass found in the last 12 months in California 215,844 jobs were posted for graduates with bachelor's degree in Data Science or Data Analytics. This represented 12% of jobs posted for graduates with bachelor's degree in the same period. The Burning Glass assessment estimated about 394,578 Data Science or Data Analytics jobs in California in 2015, increasing to 407,782 Data Science or Data Analytics jobs in 2020 and predicted to grow to 488,891 Data Science or Data Analytics jobs in 2030, which is high relative to growth in other fields. The Burning Glass assessment found that median California salary for graduates with a bachelor's degree in Data Analytics would be $103,000.

Job Types Assessment (provided in Data Science Job Types v2 12.17.2021.docx)

We have annotated the list below by boldfacing the most relevant "Data Science" and "Data Analytics" job types and key competencies associated with the proposed DSA curriculum.

Table A1: Job Types by CIP Code

Job Types by CIP Code	30.7001 - Data Science, General 207,941 jobs in California in the last 12 months	11.0199 - Computer & Information Science, General 411,899 jobs in California in the last 12 months	30.7101 – Data Analytics, General 39,469 jobs in California in the last 12 months	14.0901 – Computer Engineering, General 330,339 jobs in California in the last 12 months
Software Development Occupation Family	Software Development/Engineer **Mobile Applications Developer** Computer Scientist	Software Developer/Engineer Web Developer Software QA Engineer/Tester Mobile Applications Developer Computer Programmer Computer Scientist	Computer Programmer	Software Developer / Engineer Web Developer Software QA Engineer / Tester Mobile Applications Developer Computer Programmer Computer Scientist
Data Analysis and Mathematics Occupation Family	**Data/Data Mining Analyst** **Data Scientist** **Statistician** **Actuary**	Data/Data Mining Analyst Data Scientist	**Data/Data Mining Analyst** **Data Scientist** **Statistician**	
Network and Systems Engineering Occupation Family		Computer Systems Engineer/Architect Network Engineer/Architect Systems Analyst Cyber/Information Security Engineer/Analyst Technology Consultant		Computer Systems Engineer / Architect Network Engineer / Architect Telecommunications Engineering Specialist
Business Intelligence Occupation Family	**Business Intelligence Analyst** **Business Intelligence Architect/Developer**		**Business Intelligence Analyst**	Business Intelligence Architect / Developer
Network and Systems Support Occupation Family		Computer Support Specialist Network/Systems Administrator Network/Systems Support Specialist Webmaster/Administrator		Computer Support Specialist
Database Specialists Occupation Family	Data Engineer Database Architect Data Warehousing Specialist	Database Administrator Data Engineer Database Architect Data Warehousing Specialist		Database Administrator Data Engineer Database Architect Data Warehousing Specialist
Marketing Specialists Occupation Family	**Market Research Analyst**	Product Manager	**Market Research Analyst**	
IT Manager Occupation Family	**Chief Information Officer/Director of Information Technology**	IT Project Manager Chief Information Officer/Director of Information Technology		IT Project Manager
Front-end Application Design Occupation Family		UI/UX Designer/Developer Web Designer Video Game Designer		

Legal Support Occupation Family		Document Control/Management Specialist		
General Research Occupation Family	**Natural Science Research Manager**			
Financial Analysis Occupation Family	**Financial Quantitative Analyst**			
Engineering Manager Occupation Family				Engineering Manager
Electrical, Electronic, and Related Engineering Occupation Family				Hardware Engineer Electronics Engineer

Table A2: Key Competencies by CIP Code

Key Competencies by CIP Code	30.7001 - Data Science, General 207,941 jobs in California in the last 12 months	11.0199 - Computer & Information Science, General 411,899 jobs in California in the last 12 months	30.7101 – Data Analytics, General 39,469 jobs in California in the last 12 months	14.0901 – Computer Engineering, General 330,339 jobs in California in the last 12 months
Baseline Skills	**Communication Skills** **Teamwork / Collaboration** **Problem Solving** **Writing** **Troubleshooting** **Planning** **Research** **Creativity** **Detail-Oriented** **Microsoft Excel**	Communication Skills Teamwork / Collaboration Problem Solving Troubleshooting Writing Planning Research Creativity Detail-Oriented Organizational Skills	**Communication Skills** **Teamwork / Collaboration** **Problem Solving** **Microsoft Excel** **Research** **Detail-Oriented** **Writing** **Planning** **Creativity** **Organizational Skills**	Communication Skills Teamwork / Collaboration Problem Solving Troubleshooting Writing Planning Creativity Research Detail-Oriented Organizational Skills
Specialized Skills	Software Engineering Java **Python** **SQL** Software Development JavaScript C++ **Linux** **Git** DevOps	Software Engineering Python Software Development Java SQL JavaScript Project Management Linux Product Management C++	**SQL** **Data Analysis** **Python** **Tableau** **Data Science** **Project Management** **Machine Learning** **Business Intelligence** **Data Visualization** **Economics**	Software Engineering Java Software Development Python SQL JavaScript Project Management Linux C++ Git
Software & Programming Skills	Software Engineering Java **Python** **SQL** Software Development JavaScript C++ Linux **Git** Microsoft C#	Software Engineering Python Software Development Java SQL JavaScript Linux C++ Git Microsoft Excel	**SQL** **Python** **Microsoft Excel** **Tableau** **Data Visualization** **Microsoft Office** **Microsoft Powerpoint** Java SAP Salesforce	Software Engineering Java Software Development Python SQL JavaScript Linux C++ Git Scrum

Appendix III – Comparison of proposed DSA major and the Data Science & Computation major being proposed by faculty from School of Natural Sciences

For higher resolution, see Appendix Documents folder file "Data Science Program Comparison.xlsx".

Comparison Table between the two Data Science Proposals. Course credits are 4 semester units unless otherwise noted. Red indicates new course.

64 Units total	DS & Analytics (proposal as of January 12, 2023)		SNS: DS & Computation (Proposal as of December 14, 2022)	64 Units Total
16 Units	MATH 011: Calculus I (MATH 021 will also be accepted) MATH 012: Calculus II (MATH 022 will also be accepted) ECON 001: Intro. to Economics or COGS 001: Intro. to Cognitive Science PHIL 002: Introduction to Ethics or PHIL 003: Contemporary Moral Problems	Major Prep	MATH 011: Calculus I (MATH 021 will also be accepted) MATH 012: Calculus II (MATH 022 will also be accepted) Math 032: Prob and Stats Math 041: Matrix Analysis	16 Units
16 Units	DSA 001/MIST 060: Foundations of Data Science & Analytics DSA 002: Introduction to Computing and Logic ECON 010: Statistical Inference MIST 050: Introduction to Entrepreneurship or MIST 070: Design Thinking	Lower Division Major Requirements	Data 008: Intro to Data Science Data 011: Intro Computing and Statistical Programming for Data Scientist	8 Units
16 Units	ECON 110: Econometrics Advanced Methods A (1 of 2): MIST 130: Statistical Data Analysis and Optimization in R for Decision Support or MIST 134: Methods of Data and Network Science DSA 103/COGS XXX/PHIL 1XX: Data Ethics Advanced Methods B (1 of 3): DSA 101 Machine Learning and Natural Language Processing or DSA 102 Interactive Data Visualization or DSA 104/MIST 135: Technical Communications and Visualization	Upper Division Major Requirements	Data 100: Advanced Data Science Math 180 / Physics 127 - Machine Leaning Data 111: Intermediate Computing & Statistical Programming for Data Scientists Data 104: Ethics in Data Science	16 Units
4 Units	**Data Science and Analytics Capstone Requirement** DSA 120 Units: 2 DSA 121 units: 2		**Computational and Inferential Depth** Select three from the list (to be expanded). If the course appears on the list of "domain focus" courses, it can be counted only once to fulfill a degree requirement. Current list: Math 140, MATH 181, BIO 175 (Biostatistics), BIO 182 (Bioinformatics), BIO 184 (Object-Oriented Programming for Biologists), BIO 187 (Data Science Research in Biology), CHEM 160	12 Units
12 Units	Application Domain Elective (12 units) Complete at least three courses from the following application domain course list. Any excess courses from Advanced Methods A and B groups can also be counted towards electives **The Environment and Sustainability** MIST/ENVE 110: Hydrology and Climate ESS/BIO 113: Sustainability in the Anthropocene MIST 116/ESS 132 (3 units): Applied Climatology MIST 118/ESS 1XX: Climate Change: Science & Solutions ENGR/ESS 141: Environmental Science and Policy MIST 120: Parks and Protected Areas ENVE 130: Meteorology and Air Pollution MIST 132: Geographic Information Systems Analysis in Management PHIL 122: Bioethics ENVE 152: Remote Sensing of the Environment **Policy and Decision Making** ECON 120: Economics of the Environment and Public Policy ECON 153: Judgement and Decision Making MIST 131: Data Governance for Analytics Projects POLI 112: Public Policy: Analysis, Strategy and Impact POLI 120: Voting Behavior, Campaigns and Elections POLI 170: Theoretical Models of Politics POLI 175: Advanced Analysis of Political Data PHIL 108: Political Philosophy **Understanding and Modeling People** CSE 175/COGS125: Intro to Artificial Intelligence COGS 103: Intro to Neural Networks in Cognitive Science COGS 104: Complex Adaptive Systems COGS 105 Research Methods for Cognitive Science COGS 122: Modeling Social Behavior COGS 128: Cognitive Engineering COGS 130: Cognitive Neuroscience COGS 170: Judgement and Decision Making **Other Elective Courses** ENVE 155: Decision Analysis in Management MIST 175: Information Systems for Management MIST 1XX: User Experience MIST 190: Special Topics Course	Application Domain	At least one lower division and two upper division courses within a single major or program, preferably with some overview of how data science is applied in a specific field. **Applied Mathematics** (also requires: MATH 024, to be switched to MATH 041) MATH 140: Mathematical Methods for Optimization: Programming and optimization techniques in Matlab MATH 146: Numerical Linear Algebra: Matrix factorization and iterative methods MATH 150: Mathematical Modeling: Project-based introduction to modeling real-world systems with data MATH 170: Mathematical Biology: Modeling biological systems with programming components and data MATH 181: Stochastic Processes: Analytical and computational implementation of Markov chains and Brownian motion **Biology** (also requires: BIO 001-Contemporary Biology) BIO/ESS 113: Sustainability in the Anthropocene: Interpret, generate, and use data and multiple modes of communication BIO 141: Evolution: Data analysis and simulation in R BIO 148: Ecology: Data analysis and simulation in R **Chemistry** (also requires: CHEM 002, CHEM 008, CHEM 010) CHEM 160 (3 units): Introduction to Scientific Computing for Chemists CHEM 1XX (3 units): Data Science in Chemistry (NEW) **Environmental Systems Science** (also requires: ESS 001, BIO 001) ESS/BIO 113: Sustainability in the Anthropocene ESS/BIO 148: Ecology: Data analysis and simulation in R ESS 150: Geomorphology: Generation, analysis, and interpretation of geochemical data, and analysis of geospatial data **Physics** (also requires: MATH 023, MATH 024, PHYS 008, PHYS 009, PHYS 010) PHYS 127: Machine Learning and Statistics for Physics and Astronomy PHYS 181: Computational Physics **Materials Science and Engineering** (also requires: MATH 021, CHEM 002, PHYS 008, ENGR 045) MSE 119 (3 units): Materials Simulations MSE 114: Polymeric Materials, or other MSE courses as proposed by the MSE program **Student-Designed Custom Emphasis** Students complete between 12 and 18 units of approved course work. At least 8 of these units must be upper division courses, one of which must have a significant computing or data analysis component. Requires approval by the program Executive Committee.	12 Units

Appendix IV : Supplemental Appendix Documents – provided as addendums to the proposal in folder titled: **BA-DSA_Major_Proposal_2023_AppendixDocuments**

- DSA Undergraduate Major Program Bylaws: See **DSA_MajorProgram_Bylaws.docx** for elaboration.

- Letter of support on behalf of the UCM Library: **BS-DS-Analytics-Letter-Support-Library-2023.pdf**

- Letter of support on behalf of the proposed Gallo School department chairs: **pGalloSchool_DeptChairs_LetterofSupport_Merged.pdf**

- Letter of support on behalf of the School of Natural Sciences 'Proposal Development Team for B.S. Major or Minor in Data Science and Computing': **2023_01_30_DataScienceAndAnalytics_SupportLetter.pdf**

- Letter of support on behalf of the UCM School of Engineering Executive Committee and Dean Goel