Data Science Visions Working Group
Final Report

May, 2018
# Table of Contents

What is Data Science? ................................................................. 3  
Why Invest in Data Science at Vanderbilt? ................................. 3  
Working Group Charge and Composition .................................. 6  
Working Group Activity ............................................................. 7  
SWOT Analysis ......................................................................... 8  
Recommendation: Build a Trans-Institutional Data Science Institute .................................................. 13  
Four Year Plan for Rollout of the Institute and Educational Tracks ..................................................... 25  
Appendix A – Detailed SWOT Analysis for Education .......................................................... 28  
Appendix B – Vanderbilt Programs Related to Data Science ......................................................... 36  
Appendix C – Vanderbilt Courses Related to Data Science .......................................................... 45  
Appendix D – Vanderbilt Faculty: Data Science Experts and Users ........................................... 49  
Appendix E – Data Science Programs at Peer Institutions ......................................................... 56  
Appendix F – Select Peer Institution Case Studies .................................................................. 58  
Appendix G – Example Faculty Hires .................................................................................. 95  
Appendix H – Research IT Support for Data Science ............................................................. 97
**What is Data Science?**

Data science is an interdisciplinary field whose goal is to extract knowledge and enable discovery from complex data using a fusion of computation, mathematics, and statistics. Datasets can be as varied as maps of the universe, MRI images, human genomes, stock market transactions, historical texts, infrastructure systems or website clickstream data, among many others, making data science a natural hub for collaboration across fields. “Data science” is closely related to “big data”, which focuses on particularly large and complex datasets, but data science is more focused on the process of learning from, and making decisions based on, data rather than the size of the dataset itself, and can be applied to smaller datasets as well. Data science is revolutionizing many academic fields, while in industry it is emerging as critical training for careers in the new information-based economy.

**Why Invest in Data Science at Vanderbilt?**

Over the next decade, data science is estimated to have a significant impact across all sectors of the economy from health care to transportation, manufacturing, construction, and urban living. The power and impact of data science extends well beyond economic productivity, however, to improve individual and community health; develop smart communities that enable the efficient circulation of people, goods, and services; and enhance environmental sustainability and overall quality of life. Leveraging and amplifying the expertise of faculty, research staff, and students at Vanderbilt is critical to ensure that breakthroughs in data science are created and applied to advance discovery and learning to **tackle major societal problems in service to humanity.**

Research groups across all schools and academic disciplines at Vanderbilt are collecting increasingly large and complex data. New data analysis strategies are required to fully harness the power of these data and thus enable discovery. **Investing in data science expertise that is broadly available to, and adoptable by, Vanderbilt researchers will yield dramatic advances in academic discovery.**

**Examples of multidisciplinary collaborations**—many fostered by Trans-Institutional Programs (TIPs) awards—that can be enhanced by investments in data science at Vanderbilt include (but are not limited to) the following:

- Researchers in Engineering, Peabody, and Arts & Science are collaborating to formulate, develop, and deploy “smart” data science applications that leverage advances in machine learning to keep cities operating effectively in the face of population growth fueled by increasing commerce, facilities, and entertainment.
- Researchers in Medicine and Arts & Science are collaborating to perform big data analysis of structural biology models for vaccine development.
- Researchers at the Advanced Computing Center for Research and Education (ACCRE) are enabling faculty from diverse disciplines, such as Astronomy, Physics, and Computer Science, to develop innovative storage systems for high-energy physics experiments that help answer fundamental questions about the structure and behavior of the universe.
Researchers in Engineering and Medicine are creating advanced light microscopy technology that creates high-resolution images at the subcellular level as part of the imaging research continuum that spans from the Vanderbilt Institute for Surgery and Engineering and Vanderbilt Institute of Imaging Sciences to the Vanderbilt Center for Structural Biology, all of which open new frontiers in data science with the potential to advance early diagnosis of diseases.

Vanderbilt’s Center for Quantitative Sciences is working with 48 faculty representing 23 departments and divisions (including Physics, Chemistry, and Civil and Environmental Engineering) to streamline quantitative collaboration for improved biomedical research, including the development of personalized medicine.

Researchers in Arts & Sciences (History, French, Art History), Divinity, and Engineering are collaborating to use data science methods and tools to analyze and make accessible the physical materials of cultural heritage (such as creating searchable image databases of the medieval architecture of Asia).

Researchers in Arts & Science (Mathematics) and Engineering (Institute for Software Integrated Systems) are developing novel algorithms for detecting the source distribution of potential threats, such as pollution, chemical or biological hazards, from spatial and temporal data obtained from large numbers of cell phones or other distributed devices.

Researchers in the Law School, Arts & Sciences, and Engineering are collaborating to study the policy and legal implications of data-driven methodologies and applications of artificial intelligence.

Likewise, examples of disciplinary research that can be enhanced by investments in data science at Vanderbilt include (but again are not limited to) the following:

- Financial experts in the Law School are developing methods for detecting securities fraud by applying text analytics to 10,000 filings.
- Economic historians in Arts & Science are making new discoveries in the economic development of America by analyzing massive datasets from public agencies and connecting them to other private and public data sources.
- Peabody researchers partnering with Tennessee’s Department of Education are analyzing a massive longitudinal dataset (including data on all one million Tennessee public school students, 71,000 teachers and 1,800 principals and schools, spanning more than a decade) to answer important questions about Tennessee schools, including whether merit pay for teachers raises student achievement or whether student outcomes are improving in the state’s lowest-performing schools.
- Engineering faculty are conducting research on Bayesian statistics data analytics that integrate multiple sources of information—including historical data on weather and experts’ knowledge—to improve prediction accuracy for weather-related events, such as the impacts of a flood on inland waterways and possible recovery strategies.
- Mathematicians in Arts & Science are developing innovative methods for representing high-dimensional data by harnessing the power of sparsity and other low-dimensional geometric structures.
- Researchers in the School of Medicine are analyzing so-called “-omics” data via Vanderbilt’s DNA biorepository, BioVU, which de-identifies medical records to gain insights into the
links between genes and diseases, as well as the links between genes and patients’ response to treatment.

- Vanderbilt’s Center for Digital Humanities is digitizing historical records (including endangered sources that document the lives of slaves across the Americas) to preserve and share with the global community.

- Faculty and students from across campus are using advanced geospatial analytics and modeling tools at the Spatial Analysis Research Laboratory to pursue diverse research agendas in the social sciences, humanities, earth sciences, biological sciences, epidemiology, medicine, forensics, engineering, and education.

- Researchers in the Divinity School are using XML tree structures and semantic graph databases to model the development of networks of thousands of authors and texts in the medieval Middle East. These data are also integrated into international databases maintained by the Library of Congress and other national libraries.

- Biomedical Informatics researchers are crafting official guidance for various federal agencies in the US (e.g., Department of Health and Human Services) and abroad (e.g., European Medicines Agency) on de-identifying data derived from medical records, clinical trials, and genome sequencing projects for broad public reuse.

- Astronomers in Arts & Sciences are planning to use the next generation of astronomical surveys to study the structure of the universe and to discover habitable planets in our galaxy.

The application of data science methods, tools, and techniques is impacting every aspect of society, as shown by the ongoing projects mentioned above. Politics, social networks, marketing, criminal justice, finance, healthcare, media, education, are all being transformed by the collection and mining of large amounts of data. Human experience and judgement are often being replaced by algorithms trained on data, with implications that are poorly understood.

**Investing in the study of the impact that data science is having on society** from policy, legal, and sociological perspectives, will thus bring new areas of academic inquiry to Vanderbilt that are timely, socially relevant, and build on existing socio-legal and technological expertise.

The ability to think critically about data and the technical knowledge required to analyze data are critical skills in our information-based economy. Companies, non-profits, and government agencies are realizing the value of using their data for informed decision-making and are looking to hire people with data science skills. **Investing in both formal educational tracks and informal training in data science** will therefore give our students at all levels a competitive advantage in their future careers and will make Vanderbilt an even more attractive place for students to study.

Data collection, storage, sharing, mining, processing and visualization present challenges that span disciplines. Data science thus forms a natural hub for collaboration across fields and schools. Our compact campus and One Vanderbilt culture allows discoveries and advances to flourish. The time is now ripe to further support these endeavors and provide increased opportunities for students to study data science and for faculty to continue to push the boundary of what is possible in research, education and practice.
**Working Group Charge and Composition**

**Group Charge:** The Data Science Visions (DSV) Working Group is charged with developing a plan to leverage Vanderbilt's many assets and strengths in “big data” and “data science”, to help position the university for continued success in these areas. The working group, which is comprised of faculty from each of Vanderbilt's schools and colleges, will develop a vision for the university that will chart the path ahead. By engaging the Vanderbilt community, the working group will be better equipped to identify key thematic areas and educational programs to advance, as well as to identify opportunities for alignment and coordination across campus for greater impact.

**Hiba Baroud**, assistant professor of civil and environmental engineering  
**Andreas Berlind (co-chair)**, associate professor of physics and astronomy  
**Marisa Cannata**, research assistant professor of leadership, policy and organizations  
**Jason Grissom**, associate professor of public policy and education  
**Emily Hodges**, assistant professor of biochemistry  
**J. Kelly Holley-Bockelmann**, associate professor of physics and astronomy  
**Craig Lewis**, Madison S. Wigginton Professor of Management and professor of law  
**Bradley Malin**, professor of biomedical informatics, professor of computer science and professor of biostatistics  
**Dave Michelson**, assistant professor of the history of Christianity and assistant professor of history  
**Thomas Palmeri**, professor of psychology and professor of ophthalmology and visual sciences  
**Mariann Piano**, professor of nursing and senior associate dean for research in the School of Nursing  
**J.B. Ruhl**, David Daniels Allen Distinguished Professor of Law  
**Doug Schmidt**, Cornelius Vanderbilt Professor, professor of computer science, professor of computer engineering and associate chair of the Department of Electrical Engineering and Computer Science  
**Yu Shyr (co-chair)**, Harold L. Moses Chair in Cancer Research, professor and chair of biostatistics, professor of biomedical informatics, professor of cancer biology, professor of health policy  
**Seth Smith**, associate professor of radiology and radiological sciences, associate professor of biomedical engineering and associate professor of ophthalmology and visual sciences  
**Sonya Sterba**, associate professor of psychology and human development  
**Matt Tyska**, Cornelius Vanderbilt Professor of Cell and Developmental Biology  
**Yevgeniy Vorobeychik**, assistant professor of computer science, assistant professor of computer engineering, and assistant professor of biomedical informatics  
**Steve Wernke**, associate professor of anthropology
Working Group Activity

The DSV working group was formed in September 2017. The first phase of the committee’s work was to gather information about the research and educational programs related to data science at both Vanderbilt and its peer institutions. To organize this effort, the working group split into three sub-committees. The first sub-committee (chaired by Malin) compiled an inventory of data science research activity and organization of data science efforts at Vanderbilt’s peer institutions. The second sub-committee (chaired by Holley-Bockelmann) did the same for Vanderbilt and also investigated the needs of faculty and staff in the area of data science. The third sub-committee (chaired by Palmeri) compiled an inventory of educational opportunities in data science both at Vanderbilt and its peer institutions.

Information about peer institutions was mostly collected by studying materials that were available at institutional websites and through discussions with contacts at such institutions. Information about Vanderbilt was collected by a combination of (i) online material, (ii) the collective knowledge of committee members, (iii) surveys sent to representatives of several academic units, (iv) a series of four town hall meetings organized by the Faculty Senate, and (v) additional meetings with stakeholders from various groups on campus.

After collecting all this information, the working group then performed a SWOT (strengths, weaknesses, opportunities, threats) analysis for Vanderbilt in the area of data science. This analysis then informed the next phase of work, which was to design an institute-scale initiative in data science at Vanderbilt. To assist in this effort, committee members performed official site visits to three successful data science institutes at other universities that Vanderbilt might want to emulate. These were: the “Center for Data Science” at New York University, the “Data Science Institute” at Columbia University, and the “eScience Institute” at the University of Washington. The site visits were very helpful for the committee to understand both what aspects of these efforts work well and should be emulated and what aspects have failed and should be avoided.

The main body of this report provides a brief summary of the information gathered by the working group (via the SWOT analysis). More extensive information is included in a series of appendices at the end of the report.

- **Appendix A – Detailed SWOT Analysis for Education:** A more detailed educational SWOT analysis than the summary that is in the main body of the report.
- **Appendix B – Vanderbilt Programs Related to Data Science:** Some basic information about degree programs at Vanderbilt that are related to data science.
- **Appendix C – Vanderbilt Courses Related to Data Science:** A comprehensive list of all the courses at Vanderbilt that are related to data science.
- **Appendix D – Vanderbilt Faculty: Data Science Experts and Users:** A listing of all Vanderbilt faculty with experience or expertise in an array of data science methodologies and types of data.
- **Appendix E – Data Science Programs at Peer Institutions:** A listing of educational programs in data science at the highest ranked universities and colleges.
- **Appendix F – Select Peer Institution Case Studies:** More detailed information on the data science educational programs at eight peer institutions.
SWOT Analysis

![Faculty Proficiency in Data Science](image)

**Figure 1:** Number of Vanderbilt faculty self-identifying as users or experts in different aspects of data science. Faculty were asked to rate themselves as “expert” if they felt qualified to teach a course on the subject.

**Strengths**

- Many **groups at Vanderbilt are doing data-intensive research**, across all schools and disciplines. Over 300 faculty use some form of data science methodology in their research and over 150 self-identify as “experts” in one or more related methodologies (machine learning, visualization, statistics, simulation, data engineering, policy/ethics). In this context, “expert” means sufficiently knowledgeable to teach a course on the topic. Figure 1 provides a gross breakdown of expertise. The concentration of expertise is especially strong in the departments of Biostatistics and Biomedical Informatics, which represent major areas of strength for Vanderbilt.

- Several **world-class and unique datasets** are either hosted or heavily used by Vanderbilt researchers. Key examples are: (i) BioVU, a de-identified repository of DNA and electronic medical records from Vanderbilt University Medical Center (VUMC) patients. Data from BioVU has been integrated into the Electronic Medical Records and Genetics (eMERGE) network, which currently consists of nine academic medical centers nationwide, and for which BioVU serves as the national coordinating center. (ii) Tennessee Education Research Alliance (TERA), a longitudinal, multi-vector data set of all students and teachers in public schools in TN. (iii) Large Hadron Collider (LHC), the largest particle physics experiment in the world, for which Vanderbilt hosts data as a Tier 2 site. (iv) Sloan Digital Sky Survey (SDSS), the largest astronomical dataset in the world, with images and spectra of half a billion objects in the sky. (v) Vanderbilt Institute for Digital Learning (VIDL), which collects de-identified data from Vanderbilt’s MOOCs.
• Vanderbilt has a growing number of researchers from policy and humanities disciplines studying the impacts of data science in society. Examples of programming and curriculum they are producing include the new University Court on Ethics of AI, the Law School’s annual workshop on AI for Law/Law for AI, and the Spring 2018 symposium on the use and impacts of big data in the criminal justice system.

• Vanderbilt boasts a strong supercomputer facility, the Advanced Computing Center for Research & Education (ACCRE). This includes a hadoop cluster that was specifically built (through a TIPs grant) for big data applications. ACCRE provides solutions both for analysis and storage of large data sets.

• Vanderbilt fosters a trans-institutional culture that allows interdisciplinary endeavors like data science to thrive. Interdisciplinary work is explicitly encouraged and supported. The data science effort at Vanderbilt is being initiated with financial support from a Trans-Institutional Programs (TIPs) grant. The geographic proximity of all academic units and research centers also facilitates collaboration across disciplinary boundaries.

• Vanderbilt has a growing community of faculty and staff who have a stake in data science and are interacting with each other on a regular basis. This community has been built up over the past few years via a bottom-up approach and has accelerated this year due to the organized efforts of this working group and of the Data Science Visions TIP.

• Vanderbilt has several existing educational tracks and courses with some emphasis on data science, some of which are jointly administered by different schools. Graduate-level programs include Biostatistics (MS, PhD), Biomedical Informatics (MS, PhD), T32 NIH Training Program in Big Biomedical Data Science (PhD), Quantitative and Chemical Biology Program (PhD), and Quantitative Methods in Psychological Sciences (MEd, PhD). Undergraduate programs include the Quantitative Social Sciences major, the Scientific Computing minor, and the Quantitative Methods in Psychological Sciences minor, as well as unofficial data science emphasis in MATH, CS, EE, BME and other majors (with careful advising and course selection).

• There are several entities on campus, namely ACCRE, the University Libraries, VIDL, and the Center for Digital Humanities, that are already doing great education and outreach in data science and will be key willing partners in a trans-institutional effort.

Weaknesses

• Though Vanderbilt boasts many faculty who are experienced in some aspect of data science (e.g., statistics), there are few well-rounded expert “data scientists” in the sense that they are experts in several methodologies. Only 18 of the faculty counted in the figure on the previous page self-identify as “experts” in three or more of the categories shown. Even fewer faculty qualify as “innovators” of data science methodologies in the sense that they create and publish new methods.
• The two core disciplines of data science, computer science and statistics, are represented by relatively small programs in comparison to many of Vanderbilt’s peer institutions. Computer Science is a small program, while statistics is balkanized at Vanderbilt, with separate programs in Biostatistics, Psychological Sciences, Mathematics, Economics, and other departments.

• Faculty with access to great data have difficulty finding internal collaborators with data science expertise. This is in part because many groups across campus are disconnected from each other (especially across the VU-VUMC divide), but also because faculty with expertise in data science are stretched thin.

• Vanderbilt researchers find little support for acquiring, hosting, storing, and curating data repositories. This is especially true for groups without access to large research grants. More generally, many researchers do not have sufficient access to good research IT support.

• There are insufficient introductory courses in statistics and computer programming that are appropriate for data science at both the undergraduate and graduate level. Statistics courses are spread among different disciplines and thus tend to focus on those disciplines. Programming courses do not cover the preferred languages for data science (Python and R) and they are typically heavily oversubscribed.

• The lack of easily available formal and informal education both in basic topics, like computer programming, and in more advanced topics, like machine learning, means that faculty at Vanderbilt find it difficult to train their graduate students, postdocs, and other members of their research teams in practical data science for their research.

• Vanderbilt is a few years behind its peer and aspirational peer institutions in creating interdisciplinary data science programs. All but four of the top 20 universities in the US News & World Report rankings have some form of interdisciplinary initiative in data science and all but three have some form of graduate degree in data science.

Opportunities

• Vanderbilt has the opportunity to create a trans-institutional effort in data science that directly benefits existing research groups across campus, launches new areas of inquiry, and makes the university more attractive to prospective students by introducing new educational tracks. With sufficient investment, Vanderbilt can become a national leader in data science and compete with some of the top programs in the country, especially in some key areas like large scale medical data science, educational data science, brain science, and policy and ethics.

• With a relatively moderate investment in research staff, Vanderbilt can create an array of services to help groups across campus take their data-intensive research to the next level. These services can range from drop-in office hours and one-time think tanks, to a
longer-term incubator program. These research staff can also teach informal workshops and classes on various aspects of data science to faculty, staff, and students.

- With a higher level of investment, Vanderbilt can make **new interdisciplinary faculty hires** in data science and a variety of domains to launch new areas of research and also serve to bridge existing departments together through the common focus of data science. These new faculty can also help launch new graduate and undergraduate tracks in data science.

- A **new professional Master’s degree** in data science can help put Vanderbilt on the map in data science, lead to new courses that will be available to graduate students across the University, create incentives for industry to partner with Vanderbilt, and help make a broader data science initiative financially sustainable.

- A **new undergraduate track** in data science would be highly popular and would set Vanderbilt apart from its peers, few of which have such degrees. This would help recruit students with strong quantitative interests to Vanderbilt. Data science also provides a natural vehicle for creating meaningful and transformative undergraduate immersion experiences.

- Vanderbilt has the opportunity to help **broaden the participation of women and minorities in data science**, a field with very poor demographics. Leveraging the success of the Fisk-Vanderbilt Bridge program, Vanderbilt could become a national leader in improving the diversity of the next generation of data scientists. This would also make Vanderbilt’s educational programs especially attractive to industry, which is just now realizing the need to diversify its data science workforce.

- Nashville has a **growing and thriving local industry**, with a good mix of large companies (e.g., HCA and Nissan) and small start-ups. The landscape is thus fertile for developing strong industry partnerships in the area of data science.

**Threats**

- If Vanderbilt fails to invest in data science, academic departments in data-rich fields will increasingly fall behind those at peer institutions as data science becomes a more integral aspect of their research endeavor. Moreover, a failure to invest in new educational tracks will hurt Vanderbilt’s competitiveness for attracting students, as data science becomes more popular.

- To achieve a visible and highly ranked academic data science effort, **it is important to hire new faculty leaders** who will push the envelope both in groundbreaking applications of data science and in foundational and theoretical advances in data science itself. Building an initiative that operates solely as a service-providing unit for consumers of data science may have a substantial impact internal to Vanderbilt, but it will not lead to national and international recognition.
• The **market for faculty in data science is highly competitive** both in academia and industry. Vanderbilt will struggle to attract the best talent without a visible investment in data science.

• Hiring new faculty in data science will not automatically lead to more trans-institutional collaborations. These faculty may spend their effort completely within their home departments and not interact with each other meaningfully. The same is true for research staff. A dedicated physical space, strong leadership, and a sustained effort will be necessary to create the conditions for successful collaboration.

• Interdisciplinary faculty must be properly mentored and their trans-institutional work must be protected during the promotion and tenure process.

• The **research IT needs of faculty and staff at Vanderbilt need to be met**. Investing in data science personnel will be moot without concurrent investment in infrastructure (bandwidth, speed, storage, support).
Recommendation: Build a Trans-Institutional Data Science Institute

The recommendation of this DSV Working Group is that Vanderbilt move decisively to build a trans-institutional institute in data science. The institute should exist to promote and facilitate data-driven research in all schools and departments through interdisciplinary collaboration. Additionally, the institute should promote and facilitate education in data science, both via new undergraduate and graduate degree programs, and via informal training opportunities. For the purpose of this report, this institute will be referred to as the Vanderbilt Data Science Institute (VDSI), though of course the actual name might change.

Mission of the proposed Vanderbilt Data Science Institute (VDSI)

The priorities of the VDSI should be to (in no particular order):

- Enable research groups across campus to accelerate their data-intensive research by providing access to research staff with relevant expertise, training in data science methods, and an environment that fosters interdisciplinary collaborations.

- Promote and support foundational research in data science methodologies.

- Promote and support the study of the impact of big data on society and its institutions.

- Provide an administrative home for undergraduate and graduate degree programs in data science. Also, sponsor a wide range of informal educational opportunities.

- Build and maintain mutually beneficial connections to industry, non-profits, and local government.

- Coordinate community outreach to educate the public in matters of data.

Research activities of the VDSI

The primary goal of the VDSI should be to help researchers across campus accelerate their data-driven research by bringing people together with different backgrounds and expertise and encouraging collaboration. However, getting people to engage in interdisciplinary collaborations is challenging because everyone is busy within their own domain. Faculty tend to be overcommitted, graduate students need to complete their dissertations, and postdocs need to publish within their domains to compete for academic jobs. The solution to this problem is to provide structured activities and forums for collaboration and to ensure participation both by hiring new dedicated research staff and by providing incentives for existing members of the Vanderbilt community to be actively involved.
The working group recommends that the VDSI offer the following set of services to all members of the Vanderbilt community.

- **Regular drop-in office hours (i.e., a help desk).** Anyone should be able to come to the VDSI and consult with a data scientist. This could be for getting advice about what tool to use, for getting help with debugging code or making a data visualization or, more generally, for any issue that can be resolved quickly. This is similar in spirit to the clinics offered by the Department of Biostatistics.

- **Data think tanks.** For researchers who have data and are looking for ideas and collaborators to help them increase the sophistication of their analysis methods, the VDSI can organize brainstorming sessions that are attended by people with relevant expertise. The Data Science Visions TIP has held three such think tanks over the past year with good success. During these think tanks, the domain experts describe their data and the research questions that are interesting and then they brainstorm together with participants who have expertise in data science methods. The hope is that these think tanks will spark new research collaborations and the key to success is getting the right set of people in the room.

- **Data science incubators.** Research groups who need long-term help with a particular project can apply to an incubator program at the VDSI that can last up to several months. If selected, a member of the group (faculty, postdoc, or student) commits to spending one to two days per week at the VDSI with an assigned staff data scientist as a consultant. The team could also include Master’s students who need practical experience and undergraduates participating in an immersion experience. Over the course of several weeks the team works together intensively on the research project. When the incubation period ends, the person who was embedded in the VDSI walks away with substantial expertise that can further benefit their research group. This program is modeled after an extremely successful similar program at the University of Washington eScience Institute.

- **Community-building events.** The VDSI should organize and sponsor many events for the purpose of bringing people together and building community. Such events include data science talks, annual symposia, and roundtable discussions on specific topics related to data science. The Data Science Visions TIP has been coordinating events like this with high success.

The working group also recommends that the VDSI offer a set of extra benefits to people who are affiliated with the institute. Such people would be faculty and research staff who are hired as part of the initiative and other faculty and staff who become affiliated. For example, faculty can apply to become “data science faculty fellows” for a term of two to three years, during which they commit some level of involvement and service (e.g., attend events like data think tanks, serve on institute committees, mentor junior staff, and advise MS student capstone projects). In return, they can take advantage of the following benefits.
• **Postdoctoral fellows.** The VDSI should support a cohort of postdoctoral fellows from a variety of fields who will be fairly independent and spend most of their time at the institute. Affiliated faculty can have the opportunity to help select and advise these fellows.

• **Small research grants.** Affiliated researchers should be eligible to apply for small grants to support a variety of needs, such as help with data acquisition or data hosting, purchasing software tools, creating and supporting working groups, organizing summer institutes or workshops, supporting graduate students, etc. These would be like Discovery Grants for data science.

• **Grant-writing assistance.** Affiliated researchers can have access to research and administrative staff to help with preparing grant proposals related to data science.

• **Office space.** Affiliated researchers who commit to spending a substantial amount of time at the VDSI should have access to a shared office. An office may also be provided to Vanderbilt faculty who choose to spend time at the VDSI during a leave of absence.

• **Computing.** The VDSI should invest in some form of computing infrastructure for both storing and processing data that is free to use by all affiliated students, faculty, and staff. This infrastructure could consist of a combination of ACCRE and cloud services (e.g., Amazon Web Services (AWS) or Google Cloud).

### Educational activities of the VDSI

The DSV Working Group recommends that the VDSI serve as the administrative home for new graduate and undergraduate degree programs related to data science. Since faculty expertise in data science is spread throughout the university, the institute makes more sense as a home than any single academic department or school. In addition to formal coursework, the institute should support the teaching of informal training modules in data science skills to all members of the Vanderbilt community. Specifically, the educational components supported by the VDSI should be:

• **A professional Master of Science in Data Science degree.** This should be a standalone program within the Graduate School and will primarily target students who are looking for a stepping stone to the data science industry. The MS degree was designed and proposed by a separate task force and this committee accepts the recommendations of that task force. The MS program is also an integral part of the plan to make the VDSI financially sustainable.

• **A 4+1 option for the MS in Data Science degree.** Many Vanderbilt undergraduate students are sufficiently prepared for and would welcome a 4+1 option.

• **An undergraduate Minor and Major in Data Science.** A minor should be launched first, followed by a major if the minor is sufficiently popular. This minor can be similar in
spirit to the Scientific Computing Minor, but more focused on data science. This working group expects that such a minor program, and the associated courses that are created, will be highly popular and will also make Vanderbilt a more attractive school for many prospective students. The VDSI can also create a set of courses that will fulfill the new Digital Literacy credential for undergraduate students.

- **Undergraduate immersion experiences.** The VDSI can play the important role of matchmaker, connecting students interested in data science with on-campus research groups or off-campus industry or government entities, where they can engage in meaningful experiences.

- **Informal training modules.** There is a pressing need for informal training opportunities in basic data science skills like programming in Python or R, learning SQL and databases, using cloud computing services, etc. Furthermore, there need to be a couple different levels of instruction since some participants may have little technical preparation, while others could already be experienced. The VDSI can host these classes or it can support the University Libraries or VIDL or ACCRE, by providing VDSI research staff as instructors.

- **Public outreach.** Now more than ever, there is a need in our society for the public to be better educated in basic issues related to data. What can data tell us? When should we trust it? How should we interpret statistical studies? These are questions the public should have some familiarity with. The VDSI can play a role in filling this gap by organizing fun outreach activities (talks, school visits, and teacher training) in the local community.

**Physical Space**

Dedicated physical space that is well-designed and well-located is essential for success of the VDSI for a number of reasons. (i) The biggest risk is that faculty and staff will spend most of their time and energy within their home departments and thus not interact frequently enough to collaborate meaningfully with each other. An attractive space with collaboration specifically built in to its design can go a long way toward enticing people to spend time together. (ii) A key component of the VDSI should be a cohort of postdoctoral fellows and data science research staff from different backgrounds who serve as the “glue” to make collaboration between faculty across campus stick. The best way to nurture the cohort mentality of this team is to have them be physically located together in a collaborative space. (iii) Master’s students will need a home for studying together and attending seminars, especially given they will be paying full tuition. (iv) Attractive space will be necessary for Vanderbilt to compete with its peer institutions when hiring data science faculty or research staff, or when recruiting prospective Masters students. All the successful data science institutes at Vanderbilt’s peer institutions that this committee visited have well designed and plentiful space.

The key elements of the required space for the VDSI are

- It must be **centrally located** so that it is equally accessible from all parts of campus.
• It must be **flexible and collaborative** in its design, with a mix of use types: lots of open space, multi-use rooms, meeting rooms, and dedicated shared offices. The Wond’ry is a good model for this.

• It must be **large enough** to accommodate the permanent presence of the VDSI administrative and research staff, as well as have enough shared offices for affiliated faculty to spend time there. In addition, there must be enough space to accommodate a large number of Masters students.

The DSV Working Group anticipates that new space will need to be renovated to cover the needs of a fully scaled-up institute. In the meantime, the VDSI will need temporary space to begin to grow. For example, one possibility might be to borrow a space in the Engineering and Science building, which has several advantages: (i) it has a central location, (ii) there is plenty of space to accommodate an initial set of hires and initial MS class sizes, (iii) there are many opportunities for the VDSI to partner with the Wond’ry, for example, to help VDSI members work on developing data-driven business ideas.

**New Faculty Hires**

To make the VDSI successful and competitive with similar efforts at Vanderbilt’s peer institutions, new junior and senior faculty will have to hired. More specifically, the main reasons for hiring in data science are: (i) to build on existing areas of research strength for Vanderbilt to stake out a leadership position, (ii) to strengthen critical areas of data science that are currently weak, (iii) to teach the courses in the new proposed graduate and undergraduate programs, and (iv) to help connect schools and academic departments to the VDSI.

**Thematic areas**

The DSV Working Group anticipates that the specific thematic areas for hiring and the specific academic departments within which to hire will be determined by many factors through negotiations between department chairs, deans, and the Provost’s office. For example, one factor that should be considered is whether new hires can position Vanderbilt to be competitive for emerging big grant programs. Nevertheless, the working group expects that hires in data science will fall into the following three general categories:

• Foundational aspects of data science (e.g., computer science and statistics)
• Intensive and novel application of data science to a variety of data-rich domains
• Study of the impact of data science on society and its institutions (e.g., ethics, law, and public policy)

The working group recommends that **new hires be spread around broadly among different disciplines** rather than focused in a single area, for two primary reasons: (i) there are many possible areas where Vanderbilt could focus deeply and different segments of the University will never agree on how to prioritize them, (ii) new hires will act as ambassadors to their primary departments, which will broaden the reach of the VDSI in proportion to the number of different departments represented.
In an appendix, a number of example hires are listed as a way to illustrate the range of possible areas and disciplines that could be part of the VDSI.

- **Appendix G – Example faculty hires:** This appendix lists a number of example hires that could be part of the VDSI.

**Structure of hires**

Tenure-track faculty are appointed within schools and colleges, not institutes, and so each new hire in data science will necessarily have a primary home in an existing academic department in one of Vanderbilt’s schools or colleges. However, the DSV Working Group recommends that these faculty have a secondary affiliation with the VDSI (e.g., *Professor of Psychology and Data Science*) and that the VDSI contribute the majority of startup costs of the faculty, as well as a significant fraction of (though not exceeding 50%) of the salary. Providing initial startup costs will allow Vanderbilt to execute cluster hires in data science on a short timescale. Providing salary support in the long run is a good way to formalize the expectation that hired faculty will teach courses in data science and be active within the VDSI, in addition to their home department.

Among several risks inherent in interdisciplinary hires of junior faculty, the most problematic is failure to value and reward work outside the traditional boundaries of the faculty member’s primary domain, which can create pressure to engage less in the interdisciplinary work. Hiring and promotion procedures must be managed to safeguard against this risk. The DSV Working Group recommends that the following procedures be followed.

- **Departmental search committees** should have at least one VDSI representative, preferably from outside the department. This will help ensure that the candidates considered are a good fit for the VDSI, have the expertise required to teach data science courses, and are likely to make trans-institutional connections.

- The **expectations of the position** are made clear from the beginning to all parties. Deans, department chairs, voting faculty, and the candidates themselves all should agree on issues like the split of teaching duties between primary departments vs. data science, or the potential split between research activity in one’s primary field vs. interdisciplinary research activity.

- At the time of **promotion and tenure**, departmental promotion and tenure committees or department chairs should consult with VDSI faculty and a letter from the VDSI director should be included in the candidates’ files. This will ensure that candidates’ interdisciplinary activities and contributions to the VDSI will be recognized on the record.

- From the point of hiring, all the way to tenure, faculty should be actively mentored by senior VDSI faculty. This will help them figure out the best balance between effort spent within their field vs. in trans-institutional activities. **Proper mentoring** will also make
the VDSI feel more like a second home and will thus safeguard against the threat of hired faculty spending all their time and effort in their primary department.

- The VDSI should ideally provide a **dedicated shared office** to all faculty hired as part of the initiative. However, there should be a “use it or lose it” policy to prevent a problem of unused space.

- Over time faculty interests and trajectories can change and this freedom should be protected. For example, a senior faculty member might decide to move in a new direction within their field and away from data science. For such a faculty member who is supported by the VDSI (for example, the institute contributes a fraction of salary), there should be flexibility to amend their contract and transition them fully to their primary department.

**Affiliated Faculty**

In addition to hiring new faculty that are supported by the VDSI, the DSV Working Group anticipates that several existing Vanderbilt faculty will wish to play an active role in the institute. Affiliated faculty (especially senior faculty) will be necessary for the governance of the VDSI, as well as for mentoring new junior faculty (especially before any senior faculty are hired). The key to success will be to structure affiliations to incentivize active contribution to the institute, while a long list of affiliates that mean nothing should be avoided. The working group recommends that faculty across the university be eligible to apply for a “data science faculty fellowship”, which is a formal affiliation lasting two to three years. If selected, faculty fellows can have access to resources like mini-grants, shared office space, access to computing, etc., and in return they commit to serving on VDSI committees, and mentoring faculty and staff. Short-term rather than long-term affiliations can have several positive effects: (i) more faculty over the long run will benefit directly from the VDSI, (ii) more departments will eventually get the chance to have representatives participating in the governance of the VDSI, and (iii) the in-house expertise of the VDSI membership will constantly evolve, thus keeping it vibrant.

**Visiting Faculty**

An active visitor program can offer the benefits of having new faculty without the high cost of hiring. The DSV Working Group recommends that the DSVI create a **sabbatical program** that makes it attractive for top scholars at other universities to spend up to a year at the VDSI. Such a visiting program will also increase the visibility of data science at Vanderbilt.

**Research Staff**

The lifeblood of most successful data science institutes surveyed and visited by the working group are the postdocs and full-time research staff. These are the people who are the most
present, organize the social and educational activities, provide the glue that enables faculty to collaborate across disciplines, and actively lend their technical expertise to researchers in need of help. The DSV Working Group recommends that the VDSI follow the example of successful programs at other institutions and maintain a cohort of postdoctoral fellows and senior research staff.

- **Postdoctoral fellows** in data science typically come from a variety of academic backgrounds, but what they share in common is that they are usually pushing the envelope in terms of the data-intensive nature of their work and in terms of the advanced data science methods that they use. Most of these fellows are drawn to positions within trans-institutional institutes because they are seeking to collaborate with and learn from data science experts outside their fields. Most of these fellows want to pursue academic careers, though naturally some are lured by industry. The DSV Working Group recommends that the VDSI hire and maintain a cohort of postdoctoral data science fellows (each cohort will have a term of two to three years). The hiring search should be done across all fields and applicants should be encouraged to select a faculty advisor from within the VDSI. When selecting a cohort, care should be taken to have fellows with a mix of experience and interests. Fellows should be given complete freedom to work on whatever research they want and should be provided the support necessary for them to succeed. When hiring, preference should be given to applicants who propose to do interdisciplinary work that explicitly connects different research groups at Vanderbilt.

- **Permanent research staff** should consist of MS and PhD level people who have strong technical skills in data science methodologies and are eager to apply their skills to help researchers across campus. These scientists are not necessarily interested in following an academic path and are often more interested in building tools than writing papers. The DSV Working Group recommends that the VDSI employ a small cohort of data scientists whose main job description will be to assist and train people via the research activities outlined above (office hours, think tanks, incubator, training modules). These types of people are easily poached by lucrative industry jobs and thus some level of turnaround is expected. To improve the retention of these data scientists it is critical to (i) provide them with competitive compensation, (ii) provide them with a significant amount of freedom (for example, allow them to work on their own projects up to 30% of the time), (iii) allow and support them in developing commercial products or in launching business ideas, and (iv) have in place a framework that allows for a career path with promotions and increased responsibilities.

**Student Research Fellows and Interns**

The goal of building the VDSI is to create a vibrant data science institute to act as a watering hole for people to come and engage in meaningful interdisciplinary interaction. This environment can be enormously beneficial to students who are working on research related to data science. Moreover, having students actively involved in the VDSI will add to its vibrant culture. The DSV Working Group recommends that the VDSI formalize its connection to research students by offering **graduate student fellowships** that will support a cohort of students
doing data science-related dissertation research. With shared offices in the VDSI, these fellows will become an integral part of the culture of the institute. The students can also have access to computing and other research support provided by the VDSI. In addition, the VDSI should create a number of undergraduate summer internships, to provide highly immersive student experiences. The undergraduate interns will be placed on teams carrying out some research project and they can be jointly mentored by several team members. As an example, a project that is launched through the incubator program described previously, could consist of a faculty member and graduate student working together with a staff data scientist, a data science MS student, and an undergraduate intern.

**Administrative Structure of the VDSI**

The DSV Working Group envisions that the VDSI’s administrative home would fall under the office of the provost. Strong leadership and good governance will be critical to manage all of the research and educational programs recommended in this report. Important elements of this governance include:

- **An advisory board** that consists of senior people from both academia and industry and from both within and outside of Vanderbilt who can provide overall guidance about the mission of the VDSI and can help make connections with corporate partners.
- **A director** who is responsible for every aspect of the VDSI, such as forming committees, overseeing the budget, representing the institute in external affairs, etc. The director would report to the provost.
- **An executive director**, who runs the day-to-day operations of the VDSI and reports to the director.
- **A steering committee** that makes decisions about programs, policies, and other strategic aspects of the VDSI. The steering committee would be chaired by the director and should contain members from all relevant schools and colleges to ensure that the institute always operates in the interest of the whole Vanderbilt community.
- **Program committees**, appointed by the director and composed of VDSI-affiliated faculty and staff, who manage individual programs, such as the mini-grants, the incubator program, the undergraduate internships, etc.
- **Administrative staff**, who assist with operations, budgets, space issues, IT, and industry outreach and report to the executive director.

In addition to this governance, the new educational degree programs need to have their own separate administrative structure. Most importantly, each track needs a faculty director and a faculty program committee. Specifically, a **director of graduate studies** would manage the MS program and a **director of undergraduate studies** would manage the undergraduate minor and major. The MS program in particular will need significant staff support given its anticipated size. The working group anticipates that the directors of educational programs will work closely with the director of the VDSI to align the research and educational missions of the institute.
Computing Infrastructure of VDSI

The hardware and technical support needed to store, manage, and process large data sets is expensive and not equally or sufficiently available to all researchers. This is especially true for people in the social sciences and humanities who do not typically have access to large grants. The DSV Working Group recommends that the VDSI invest in some level of computing infrastructure (both for storing and analyzing data) and make it freely available to its affiliated faculty, staff and students. The exact form of infrastructure will depend on the specific needs, but will likely consist of a hybrid model that includes both some dedicated hardware at ACCRE, and some support for the use of cloud services (e.g., AWS and Google Cloud).

Financial Sustainability of the VDSI

Launching an institute of the scope needed to position Vanderbilt as a leader in data science will require substantial initial investment, mostly in the form of startup packages for new faculty. This investment must come from the University, potentially via a fundraising campaign. However, the goal of the VDSI should be to become financially sustainable as quickly as possible. Vanderbilt’s peer institutions with trans-institutional institutes like the one described in this report use four main sources of revenue to achieve sustainability.

- **Tuition revenue from a MS program.** This is by far the largest and most reliable form of revenue used by Vanderbilt’s peers. Data Science MS programs are immensely popular (some receive 1,000 to 2,000 applications) and students find lucrative jobs upon graduation so they are willing to invest in the degree.
- **Overhead from grants.** Many data science institutes encourage faculty to submit grants through the institute and then collect a portion of the indirect costs. In most cases of peer institutions studied by this working group, this tool generates only modest revenue. Moreover, it comes with the added cost of creating tension between the institutes and the colleges/schools that are competing for overhead. This is not a good way to generate goodwill and cooperation between the VDSI and the schools and colleges.
- **Industry outreach.** Some of the established data science institutes successfully court industry partners, who then pay fees to become formally affiliated. Participating companies are mostly interested in having preferential access to students who graduate with technical skills. They visit periodically, talk to students, and participate in job fairs. Other data science institutes have not been successful at this. In short, industry support is possible, but takes substantial effort to achieve.
- **Grant programs.** There have been some large grant opportunities in the past few years that have funded large efforts in data science, such as the Moore-Sloan Data Science Environments or the NSF Big Data Regional Innovation Hubs, and the NIH Big Data Centers of Excellence. These grants, however, are to scarce to offer sustainability, and so have been used primarily to launch rather than to sustain institutes.

The only way to make the VDSI sustainable is through the MS program. The DSV Working Group therefore recommends that the revenues from the MS program first be used to cover the costs of the program itself and that the remaining revenue be used to support the VDSI. By
growing the MS program to the right size (with 20% allocated to scholarships in order to ensure diversity), the VDSI could become financially sustainable. The working group also recommends that the VDSI cultivate relationships with local industry as a possible secondary source of revenue.

**Target Size of Institute**

The size of the VDSI (number of faculty and staff hired) will of course depend on the level of initial investment by the university. The DSV Working Group has estimated a rough target size based on the following factors: (i) there are enough new faculty hires to have a big impact both by building strength in core areas and by achieving diversity of fields, (ii) there are enough new faculty to teach the courses that will be developed for the MS and undergraduate minor and major tracks, and (iii) there are enough postdoctoral fellows and research staff to establish cohorts of critical mass and to meet the needs of researchers across Vanderbilt. Based on these factors, the VDSI should aim to reach a size of roughly 20 new junior faculty, four new senior faculty (where the faculty are divided roughly equally between departments in VUMC and the School of Medicine, and in all the other Schools and Colleges), one visiting faculty, four senior data scientists, eight postdoctoral fellows, eight graduate fellows, and eight undergraduate interns. The administrative team of the VDSI will need to include one faculty director, one executive director, one IT staff, and two administrative support staff. In addition, the proposed MS in Data Science program will require one faculty director, one program coordinator, and one administrative support staff.

The VDSI Working Group has estimated that an institute of this size can become financially sustainable in the time it will take to grow the MS program to a size of 60 students per class (allowing for the equivalent of full scholarships for 12 of these students), with some additional revenue from an industry affiliates program. This sustainability plan includes the cost of all staff, programs, and faculty salary support. A modest initial investment is required to cover the operational costs of the institute (staff and programs) as it ramps up and before it achieves sustainability. A much larger investment is needed to cover faculty startup packages and the cost of renovating space for the institute. In the event of a significantly lower initial investment, the VDSI can scale back to fewer new faculty than what is proposed here; however, an absolute minimum of six full-time faculty are needed to teach the MS courses, assuming their whole teaching load is in the MS program. A scaled-back institute can still achieve the goals of assisting researchers with their data needs and providing informal training to the university community. Such an institute would fill an important service role, but it would not become a source of new scholarship and it would not make Vanderbilt competitive with its peer institutions in the area of data science.
The final appendix of this report discusses how the parallel efforts to improve research IT at Vanderbilt can both benefit and benefit from the proposed data science initiative.

- **Appendix H: Research IT Support for Data Science:** This appendix summarizes the key findings of the Research IT Working Group and outlines the synergies between the proposed research IT services and the VDSI.
The DSV Working Group recommends that the process for launching the VDSI and its associated educational programs begin immediately so that the benefits to Vanderbilt students and researchers are felt as soon as possible. At the same time, the rollout should proceed in a fiscally responsible way. This means that the top priority is to get the MS program up and running quickly, with a target launch date of Fall 2019. The first class should be small (up to 20) and can expand over the next few years to the target size of 60 as kinks are worked out and as the program demonstrates success in providing all students with meaningful practical experiences and in placing students well after graduation. What follows is a high-level outline of a four-year plan to scale the VDSI up to its target size, with important steps to be taken each year.

Spring 2018

- **MS Program:**
  - Appoint a director of graduate studies (DGS) for the MS program. Much planning needs to be completed quickly if the MS program is to be launched in Fall 2019 and the DGS needs to start working on a detailed plan as soon as possible.
  - Appoint a graduate program committee (GPC) to assist the DGS with planning.

- **VDSI:**
  - Appoint an interim director for the institute. The working group anticipates that at some point in the future the VDSI may search for an external director, but until then, there needs to be an interim director to lead the rollout. The interim director and the DGS will work closely together.
  - Appoint a steering committee to help the interim director with planning.

Fiscal year 2019

- **MS Program:**
  - Hire program coordinator and assistant to DGS for the MS program.
  - Over the summer, finalize the structure of the program so it can be properly advertised in the fall.
  - Address all administrative hurdles inherent in starting a new degree program.
  - Advertise the program and solicit applications for the two-year track. Select the first class.

- **Undergraduate Minor:**
  - Appoint a director of undergraduate studies (DUS) for the data science minor program and start planning the structure of the program.

- **VDSI:**
  - Orchestrate search for a first cohort of junior hires. Hire a minimum of six new faculty to begin to teach when the MS program launches.
o Solicit applications for data science faculty fellowships and select a first cohort. Form various program committees.

o Allocate temporary space for the VDSI.

o Begin the study and design process for renovating more permanent space.

o Search for and hire two senior data scientists so that they can support and lead research across campus.

o Search for and hire a small cohort of three to four postdoctoral fellows.

o Hire an administrative assistant for the interim director.

o Following the hiring of research staff, rollout select institute services (office hours, think tanks, incubator, training modules, mini-grants).

**Fiscal Year 2020**

- **MS Program:**
  o Fall of 2019: launch the MS program with up to 20 students.
  o Solicit applications for the next class, this time opening it up to 4+1 applicants.

- **Undergraduate Minor:**
  o Launch data science minor.

- **VDSI:**
  o Orchestrate search for a second cluster of six junior hires.
  o Search for and hire an executive director.
  o Continue providing services of the institute.
  o Begin renovation of permanent space.
  o Invest in computing infrastructure.

**Fiscal Year 2021**

- **MS Program:**
  o Fall of 2020: enroll the second class with up to 40 students, including some 4+1 students.

- **VDSI:**
  o Continue renovation of permanent space.
  o Continue providing services of the institute.

**Fiscal Year 2022**

- **MS Program:**
  o Fall of 2021: enroll the third class with up to 60 students.
- **Undergraduate Major:**
  - Launch data science major.

- **VDSI:**
  - Move into permanent space.
  - Hire remaining research staff.
  - Orchestrate search for a third cluster of six junior hires.
APPENDIX A: Detailed SWOT Analysis for Education

Strengths

- **Vanderbilt has significant expertise in many of the component areas of data science**
  - The DSV working group considered the component areas of “data science” to be an interdisciplinary combination of computer programming, simulation, mathematics, statistics, visualization, machine learning, and domain knowledge.
  - This expertise is reflected in the current programs at Vanderbilt related to data science (details in Appendix B): Biostatistics (MS, PhD) (SoM), Biomedical Informatics (MS, PhD) (SoM), T32 in Big Biomedical Data Science (SoM), Quantitative and Chemical Biology Program (PhD) (A&S, SoM), quantitative area in Psychological Sciences (Undergraduate Minor, MA, PhD) (A&S, Peabody), Econometric emphasis areas, proposed (BA/BS) Quantitative Social Sciences major (A&S, Peabody), Scientific Computing undergraduate minor (A&S, Eng), unofficial data science emphasis in MATH, CS, EE, BME and other majors (with careful advising and course selection), as well as a Big Data Science Intellectual Neighborhood in the School of Engineering.
  - This expertise is reflected in the current course offerings across numerous colleges and departments at Vanderbilt (details in Appendix C) in introductory and advanced programming, introductory and advanced statistics, mathematical fundamentals, visualization, big data and informatics, machine learning and deep learning, computation and simulation.
  - This expertise is also reflected in the research interests of Vanderbilt faculty, who currently make use of data science methods, would like to use these methods but do not have the requisite expertise, or see the importance of training the next generation to use these methods to solve next-generation problems.

- **Vanderbilt has unique strengths to create world-renowned education programs in data science**
  - Vanderbilt has numerous highly visible and highly ranked educational and research programs that can both support and capitalize on the creation of data science programs and courses. Some of these programs already have outstanding concentrations related to data science, others could be significantly enhanced by the creation of data science education programs and courses, research support, and infrastructure.

- **Vanderbilt fosters a trans-institutional culture that will allow a multidisciplinary educational endeavor like data science to thrive**
  - One Vanderbilt and the investment in Trans-Institutional Programs (TIPS) aims to create new trans-institutional endeavors. Data science is a canonical example of the kind of multidisciplinary effort One Vanderbilt and TIPs aim to support. Case in point: the Data Science Visions effort is being initiated with financial support from a ViA TIPs grant.
  - Vanderbilt already has examples of graduate and undergraduate education programs that cross traditional disciplinary and college boundaries. At the graduate level, the Interdisciplinary Neuroscience Program includes faculty and students from the College of Arts and Science, Peabody College, Engineering, the Medical School, Nursing, and
the Law School. At the undergraduate level, the Scientific Computing minor was the first cross-college undergraduate minor and includes faculty from fifteen departments in three colleges.

- University Courses provide a mechanism to new and creative trans-institutional learning; one relevant current example is the course Data Science Methods for Smart City Applications.
- Vanderbilt supports trans-institutional endeavors that can synergize well with new efforts in data science. Salient examples include the Advanced Computing Center for Research and Education (ACCRE), the Vanderbilt Institute for Digital Learning (VIDL), and the University Libraries. All variously provide hardware, software, and storage resources and training workshops on topics relevant to education and research in data science.
- Vanderbilt has a collaborative culture, social networks, and geographic accessibility that increases the likelihood of success of trans-institutional, multidisciplinary endeavors like data science.

Weaknesses

- **Statistics – a core component of data science – is currently balkanized at Vanderbilt.**
  - Programs and coursework exist in Biostatistics, Economics, Mathematics, Psychological Sciences, and other departments. Unlike many peer research universities and liberal arts colleges, Vanderbilt does not have a department of Statistics or a department of Mathematics and Statistics.
  - As a result, many courses currently offered in statistics are focused on statistics for particular disciplines. These courses may require prerequisite courses in a particular discipline, may use examples solely from a particular discipline, may be closed to non-majors entirely, or may be oversubscribed.
  - Although a statistics-related option is being developed at the undergraduate level – the Quantitative Social Science Major – it does not emphasize data science and requires specialization in a specific social science field after introductory coursework.
  - No general statistics options exist at the graduate level.

- **Current coursework in statistics is often not ideal for data science.**
  - As noted above, many introductory statistics courses offered outside of Mathematics are discipline-focused.
  - At the undergraduate level, the MATH 1010/1011 sequence is a basic introduction not appropriate to students in the sciences or engineering, therefore it cannot be a basis for data science. The MATH 2810/2820 sequence requires as prerequisites the full calculus sequence, which is not always taken by students in some science and social science disciplines, therefore it cannot be a basis for data science. Moreover, the prerequisites for MATH 2810/2820 would force the statistics sequence into the junior year, making it too late as a core for any undergraduate data science major to be viable.
  - Elements of linear algebra and discrete mathematics are also core topics relevant to data science. At the undergraduate level, linear algebra courses like MATH 2400 or
2410 and discrete mathematics courses like MATH 3700 require as pre/co-requisites the full calculus sequence. And at the graduate level, no linear algebra or discrete mathematics courses for non-mathematics majors are offered. A course on mathematical concepts relevant to data science could be developed to cover topics such as these for non-mathematics majors interested in data science.

- Outside of particular content disciplines (like Biostatistics, Psychological Science, Economics), no current statistics course sequence exists to support a Master’s program, graduate certificate, or Doctorate in data science.

- Traditional approaches to statistics focus on descriptive and inferential statistics with small numbers of explanatory variables – think paired t-tests, single-variable regression, one-way ANOVA. A data science approach to statistics might well introduce multidimensional data far earlier in the pedagogy.

- Traditional approaches to statistics focus on mathematics and proof, or on using packages like Minitab (at the undergraduate level) or SPSS or SAS (at more advanced levels). Data science requires deep knowledge of programming-based approaches using languages like R or Python; indeed, most job advertisements in data science require these programming skills. A data science approach to statistics would introduce programming as a way to understand and solve problems from day one.

- **There are too few opportunities to learn computer programming at Vanderbilt and the opportunities that exist are not ideal for data science.**
  - At the undergraduate level, CS 1101 uses Java, CS 1103 uses Matlab. Arguably, Python would be a more appropriate introductory language for students interested in data science; the undergraduate minor in Scientific Computing requires students to learn Python. In data science, Python is the ideal programming language to learn methods for general programming and simulation, R is an appropriate programming language to learn statistical analyses.
  - CS 1101 is billed as an “intensive introduction”, which may be perfectly appropriate for CS majors, but may scare off some students who otherwise would benefit from and/or have an aptitude for programming, especially in applications to data science.
  - CS 1103 is geared towards engineering students, who need to learn Matlab for their advanced engineering coursework, and is often closed to non-engineering students. Matlab is useful for some work, but is not as widely used as Python or R in data science.
  - Outside of one recently-created course in Matlab and a few discipline-specific programming courses, there are no general computer programming courses for graduate students.
  - Additional routes to learning computer programming for undergraduates and graduate students are needed. CS has recently begun to solve this by creating a new CS 1000 course for non-majors as an entry to programming. Additional entries should be created, such as an introductory CS course for students interested in the social sciences, with a concentration in data science in particular.
  - In addition to introductory programming courses, there should be more advanced programming courses for non-CS majors at both the undergraduate and graduate levels. The creation of the Scientific Computing minor did create a new 2nd-semester
programming course (CS 2204) for undergraduates; there should be graduate courses offered for non-CS majors.

- Demand for intermediate courses like CS 2201 and 2204 is growing, with cause for concern that students will be bumped from these courses because of over-enrollment.
- Vanderbilt should commit the requisite teaching staff to introductory and intermediate computer programming courses so that students will not be bumped from these courses in the future. These courses will be core to any data science program and students in those programs must be able to enroll in them.

- **Current computer programming courses are often oversubscribed and often too large.**
  - Sections of CS 1101 often reach capacity and CS 1103 is often closed to non-majors.
  - The sections of CS 1101 and 1103 that are offered are too large for an introductory programming course, with 100 or more students. While some material may be taught effectively in a large lecture format, learning programming well requires more opportunities for student interactions with faculty and other instructors.

- **There are too few formal opportunities to learn new programming languages outside of introductory courses.**
  - Students learn Java in CS 1101, Matlab in CS 1103, C++ in CS 2201, Python in CS 2204.
  - If a student who knows Java wants to learn Matlab, or if a student who knows Python wants to learn C++, they need to take a full course or learn on their own. This is not ideal since both CS 1101 and 1103 are introductory courses and since both CS 2201 and 2204 are both data structures courses, covering many of the same concepts.
  - There should be other avenues for learning a new programming language (Java, Matlab, C++, Python, R, Julia) – at the undergraduate and graduate levels – for students who know how to program. These could, for example, take the form of 1-credit hour modules taught at the beginning of every semester.
  - The lack of such modules has an impact on how current courses are taught. For example, a faculty member who wants students to use Matlab either needs to restrict enrollment to students who took CS 1103, spend several weeks at the start of the course teaching Matlab instead of core content, or ask that students learn Matlab on their own. If there were a module offered, the faculty member could instead require/recommend students to take that at the start of the semester.
  - These could be online MOOC-style offerings, taught by Vanderbilt faculty or as a carefully curated selection taught from people outside of Vanderbilt.

- **Current Vanderbilt programs (undergraduate and graduate) often struggle to increase the quantitative and/or computational skills of their students.**
  - Because there are too few routes to learning statistics and computer programming at the undergraduate and graduate levels, programs are either left with having students graduate with missing knowledge and skills or devote scarce teaching resources to teaching these skills.
• Despite the large selection of course offerings, there are notable gaps in course offerings for core and advanced data science topics.
  o There are courses in machine learning, deep learning and neural networks, visualization, exploratory data analysis, simulation methods, high performance computing and big data.
  o These are often taught too infrequently or are taught with too much of a tight disciplinary focus to be viable core courses for a data science program.

• Vanderbilt does not yet have a data science program.
  o While data science is a relatively new interdisciplinary field, most of Vanderbilt’s peer institutions are well ahead of Vanderbilt in implementing new data science programs, especially at the graduate level.

Opportunities

• Data science provides an opportunity to create a set of truly interdisciplinary, trans-institutional education programs at Vanderbilt.
  o There is no single, natural home for data science at Vanderbilt. Expertise in the components of data science is spread across many departments and colleges.
  o Successful data science programs at Vanderbilt will require leadership, oversight, and teaching from across the university.
  o Creating programs in data science can create new opportunities to promote stronger interdisciplinary research at Vanderbilt.

• Data science offers opportunities for new interdisciplinary and trans-institutional faculty hires to recruit individuals to teach data science courses.
  o With its combination of programming and simulation, mathematics and statistics, as applied to important domain problems, data science requires faculty with training that spans traditional disciplinary boundaries.
  o Imagine a new faculty hire in an area like digital humanities, or one who combines neuroscience methods with big data techniques, or one who combines machine learning and astrophysics. These individuals could not only teach courses in their core discipline but could teach more general data science courses to a broader audience of students.
  o The creation of new data science courses creates an opportunity to see faculty expertise in the Medical School, most notably in Biostatistics and Biomedical Informatics, impact graduate as well as undergraduate education more broadly at Vanderbilt. One challenge, however, is creating incentives and effort allocations to make this teaching equitable with those of faculty in other units on campus.

• Creating new courses to support data science programs will have significant impacts on other programs and disciplines at Vanderbilt outside of data science.
  o Creating new course pathways to learning statistics, programming, simulation, and modeling can only help to improve the many undergraduate and graduate programs at
Vanderbilt wishing to increase the quantitative and computational sophistication of their students.

- Creating workshops and modules related to data science will help students, staff, and faculty develop new quantitative, statistical, and computational skills.

- **Creating programs in data science at Vanderbilt will help attract the best and brightest to Vanderbilt’s undergraduate and graduate programs.**
  - On a weekly basis, major newspapers and magazines have articles about data science as a growing discipline with numerous career opportunities.
  - Many students – and their parents – will likely pay close attention to whether universities have data science programs, and pay attention to its offerings in the components of data science, such as statistics, programming, and simulation.
  - A failure to create a world-class data science program at Vanderbilt could be deemed a threat to the institution and its standing.
  - Even if data science moves from being the hot new thing to a more established and more common discipline, any courses and programs created to support data science are core to most social science, science, and engineering disciplines and are of growing interest to the humanities and other disciplines.

- **In addition to creating graduate programs in data science, Vanderbilt has a unique opportunity to create an undergraduate program in data science given its long-standing excellence in undergraduate education.**
  - While most peer research universities have, or will have, Master’s degrees in data science, few have undergraduate degrees. Some liberal arts colleges are creating some kind of support for data science.
  - Creating a strong undergraduate presence in data science would help recruit students with strong computational and quantitative interests to Vanderbilt who might otherwise gravitate to universities with traditional strengths in those areas, like Stanford and Carnegie Mellon.
  - The welcome recent creation of cross-listings of courses as both undergraduate and graduate can go a long way to supporting both undergraduate and graduate data science programs at Vanderbilt.

- **Data science provides a natural vehicle for creating meaningful and transformative undergraduate immersion experiences.**
  - With its natural interdisciplinary nature, bridging mathematics and statistics with computational in a domain area, undergraduates in data science can work on an array of problems.
  - These include basic and applied research, business, government, non-profits, and political organizations.

- **There are too few women and underrepresented minorities in quantitative and computational disciplines, and data science programs could help to increase those numbers.**
  - Data science, being a highly interdisciplinary field, has the potential to offer underrepresented groups a route to combining statistical, mathematical, and
computational work with meaningful domain applications of these tools. This kind of program could help attract more women and underrepresented minorities in comparison to more traditional fields with quantitative and computational foci (computer science, engineering, mathematics), which often struggle to maintain equity.

- Data science is not only an interesting course of study, but also, at least currently, a lucrative field, making it potentially attractive to a wide array of students.

- Nashville is a rapidly growing city with a thriving economy and significant growth expected in data-science-related industries, like health care, technology, distribution, manufacturing, music and entertainment. Data science provides unique opportunities for collaboration between Vanderbilt and local industries.
  - Data science leaders in local industries can provide expertise to guest lecture in courses, provide expert guidance on education programs, and provide internship opportunities and future employment to Vanderbilt graduates.
  - A strong data science presence at Vanderbilt could lure new tech industries to Nashville.

**Threats**

- While creating a true trans-institutional set of programs in data science offers unique opportunities, it also carries significant potential threats if leadership in data science are not given faculty resources to guarantee that courses are taught and programs are administered.
  - A significant challenge to any trans-institutional program is that faculty time is typically controlled by department chairs and college deans, not the programs themselves. These institutional leaders could see a program like data science as outside of their core mission and not give up faculty time to teach courses or administer programs.
  - Core faculty need to be committed long-term to making a program like data science succeed. Courses that are created should not be taught by adjunct faculty or be shuttled around to faculty with short-term appointments. Faculty interested in committing to developing and teaching courses in data science should be guaranteed that they will be able to teach those courses on a regular basis.
  - Teaching these courses should also not be considered an add-on load for faculty.
  - The university needs to be creative in developing a viable solution. Leadership in data science should be committed faculty FTEs and/or provided resources to buy faculty time.
  - Faculty might receive an official joint appointment, say as “Professor of Data Science” to acknowledge their efforts, an approach that has been used by some peer institutions.

- The trans-institutional nature of data science would be threatened by a move to compartmentalize data science within an existing unit of the university.
  - The university needs to be creative in development a governing structure for data science to preserve and enhance its trans-institutional nature.
A failure to make a significant investment in data science can threaten to make it difficult to recruit new faculty and build a world-class program of research and education.
  o With so many top peer institutions building programs in data science all at the same time, there will be significant competition for recruiting top faculty.
  o A concerted and significant effort, like the one Vanderbilt made when it began building the Institute for Imaging Science (VUIIS) two decades ago, should be considered.

Some college-specific rules may need to be re-examined when it comes to the development of trans-institutional programs like data science.
  o Some college rules are hostile to trans-institutional programs. Most notably, the College of Arts and Science requires its undergraduates to have 102 credit hours of coursework in the college, which poses a challenge to majors/minors with significant coursework outside of the College.
APPENDIX B: Vanderbilt Programs Related to Data Science

**PhD Programs**

**Biomedical Informatics (Medicine), PhD**

Web Site: [https://www.vumc.org/dbmi/research-ms-and-PhD-program](https://www.vumc.org/dbmi/research-ms-and-PhD-program)

Contact: Cynthia Gadd

Description: Vanderbilt DBMI’s MS and PhD in Biomedical Informatics is an internationally recognized research program, in which students actively engage in high quality research mentored by DBMI’s 60+ faculty. As a result, DBMI students have achieved outstanding publication records, as well as numerous recognitions from national organizations. Candidates with prior training in the computational and biosciences are preferred. Funding opportunities include our National Library of Medicine Biomedical Informatics Training Grant and postdoctoral research fellowships, as well as the VA Postdoctoral Fellowship in Medical Informatics & Quality Improvement.

Target Student: a student with interest in Biomedical Informatics

Brief Curriculum Outline: courses in bio and medical informatics foundations, machine learning, data analysis, and biology

**Biomedical Informatics and Data Science (BIDS) (Medicine), PhD T32**

Web Site: [https://www.vumc.org/dbmi/vanderbilt-big-biomedical-data-science-bids-program](https://www.vumc.org/dbmi/vanderbilt-big-biomedical-data-science-bids-program)

Contact: Bradley Malin

Description: A confluence of technical, analytical, and policy advancements have thrust the biomedical community into the big data science age. Vanderbilt University is right at the center of it all and is the place where the next-generation of biomedical data science investigators and practitioners will be educated and endowed with the skillset to innovate new technologies and analytic strategies to support basic scientific advances. The Vanderbilt Big Biomedical Data Science (BIDS) Training Program lays a foundation in, and emphasizes the symbiotic relationship between, Biomedical informatics, Computer Science, and Biostatistics. The BIDS program provides PhD students with access to a diverse array of real big biomedical data sets, software tools, and applications at Vanderbilt (and interdisciplinary collaborations). BIDS also integrate courses and faculty from across the institution to ensure that students are well-versed in the foundational competencies of computation, statistics, and biomedical science that are necessary to achieve reproducible success in this field.

The overarching objective of the Vanderbilt BIDS Training Program is to thoroughly prepare the future leaders of the biomedical community focused on infrastructure, software tool development, and big data analytics. The BIDS program aims to achieve this goal through rigorous classroom and research training, as well as career development, for predoctoral trainees who will matriculate into the new Data Science Track of the existing Vanderbilt BMI PhD program. For more information see the Curriculum Overview of the Data Science Track.

The BIDS Faculty Mentors cover the foundational areas of biomedical informatics, biostatistics, and computer science (with a focus on databases and machine learning). Yet, perhaps more important to the success of BIDS is the fact that, the faculty mentors also cover a range of biomedical scientific disciplines (e.g., biochemistry, cancer biology, and genetics) and clinical application domains, (e.g., anesthesiology, internal medicine, and oncology), that will provide clear illustrations of types of data and environments big data technologies need to
support. It should further be noted that an overwhelming majority of the faculty mentors are involved in large team-based scientific initiatives (e.g., eMERGE, PGRN, regional health information exchange implementation and evaluation) and have experience in co-mentoring of students (often across schools of the university).

Target Student: PhD student interested in Biomedical Informatics and Data Science

Brief Curriculum Outline: a mix of required and elective courses in DBMI, Computer Science, Biostatistics, and Nursing/Biology

**Biostatistics (Medicine), PhD**

Web Site: [http://www.vanderbilt.edu/biostatistics-graduate/current-program/phd-program/1756-2](http://www.vanderbilt.edu/biostatistics-graduate/current-program/phd-program/1756-2)

Contact: Jeffrey Blume

Description: Biostatisticians are interdisciplinary scientists who specialize in the interpretation of data as scientific evidence. They use mathematical frameworks to account for uncertainty, extract knowledge from data, and generalize results to target populations. The discipline of Statistics laid the foundations for how science learns from data. This process, combined with recent advances in computing, has blossomed into the emerging field of “Data Science”. Vanderbilt’s Biostatistics program seeks to train the next generation of statistical leaders who are dedicated to the biomedical sciences and public health. They will design and conduct experiments, analyze data and interpret results, refine data capture techniques in mobile devices and images, and develop new statistical methods to support and encourage cutting edge scientific discoveries. Prediction algorithms developed by our students are already used for routine medical care here at Vanderbilt University and at other hospitals across the country. Many of these algorithms, and their novel underlying mathematical tools, will soon be available to you on your mobile device.

Target Student: PhD student interested in statistics and clinical applications

Brief Curriculum Outline: foundational and applied statistics courses

**Biostatistics (Medicine), Minor for PhDs in other disciplines**

Web Site: [https://www.vanderbilt.edu/biostatistics-graduate/current-program/biostatistics-minor/](https://www.vanderbilt.edu/biostatistics-graduate/current-program/biostatistics-minor/)

Contact: Jeffrey Blume

Description: The Biostatistics minor is a departmental mechanism to encourage quantitatively oriented students to broaden their statistics background while receiving official acknowledgement from the Department of Biostatistics. The Graduate Program in the Department of Biostatistics will certify when students have completed their requirements, allowing those students to list this designation on their Curriculum Vita.

Target Student: Graduate students who are currently enrolled in a PhD program at Vanderbilt University, and who are in good standing in their respective program, are eligible to participate in the Biostatistics minor program.

Brief Curriculum Outline: foundational and applied statistics courses

**Quantitative and Chemical Biology Program (Medicine), PhD**

Web Site: [https://medschool.vanderbilt.edu/qcb/](https://medschool.vanderbilt.edu/qcb/)

Contact: Carolyn Berry

Description: The Quantitative and Chemical Biology program is a PhD track, multidisciplinary program introducing elements of biology to students who wish to pursue a doctoral degree at the interface of the chemical, physical, and biological sciences. The curriculum prepares students for
research careers in any area of biomedical research. Many students choose to work in interdisciplinary fields such as chemical biology, structural biology, imaging sciences, molecular and cellular biophysics, or systems biology. Participating programs and departments include Biochemistry, Biological Sciences, Cancer Biology, Cell & Developmental Biology, Chemical & Physical Biology, Chemistry, Human Genetics, Mathematics, Microbe-Host Interactions, Molecular Pathology & Immunology, Molecular Physiology & Biophysics, Neuroscience, Pharmacology, Physics & Astronomy, Biomedical Engineering, and Chemical & Biomolecular Engineering.

**Target Student:** Previous didactic training in the biological sciences is not required for entry into the QCB program.

**Brief Curriculum Outline:** https://medschool.vanderbilt.edu/qcb/program-overview

**Computer Science (Engineering), PhD**  
**Web Site:** https://engineering.vanderbilt.edu/eecs/Graduate/Programs.php  
**Contact:** Akos Ledecki  
**Description:** The computer science program provides opportunities for students to join vibrant collaborative and interdisciplinary efforts in artificial intelligence, computer animation and virtual environments, cyber-physical systems, distributed real-time and embedded middleware, human-systems integration, image processing, intelligent learning environments, model-integrated computing, robotics, software engineering and trustworthy computing. Students can explore a broad range of research areas in computer science, as well as interdisciplinary research opportunities in areas related to the School of Engineering's strategic directions in health care and medicine, security, energy and natural resources, and Data Science.  
**Target Student:** anyone with an aptitude for STEM.  
**Brief Curriculum Outline:** Core courses in computing, data structures, and algorithms, and elective/depth courses (AI, Databases, Big Data, Machine Learning) that allow a Data Science specialization

**Economics, Econometrics Specialization (A&S), PhD**  
**Web Site:** https://as.vanderbilt.edu/econ/graduate/  
**Contact:** Mattias Polborn  
**Target Student:** Most students entering our Ph.D. program have majored in economics as undergraduates, and some have a Master's degree in economics. Extensive training in economics is helpful but not sufficient preparation for the Ph.D. program. It is essential that each entering student has taken at least one year of calculus, one semester of statistics, and intermediate level courses in microeconomic theory and macroeconomic theory. Courses in linear algebra differential equations, real analysis, and econometrics are also strongly recommended. Any deficiencies in these areas should be made up before beginning graduate work.  
**Brief Curriculum Outline:** Microeconomics core (8100, 8110, 8120), macroeconomics core (8200, 8210, 8220), statistics and econometrics core (8300 Probability and Statistics, 8310
Econometrics I, 8320 Econometrics II), economic history core (8400), Mathematics core (8000 Math for Economists). Econometrics specialization courses include: ECON 9310 Time Series Econometrics; ECON 9320 Nonparametric and Semi-parametric Econometrics; ECON 9330 Microeconometrics

**Mathematics (A&S), PhD**

**Web Site:** [https://as.vanderbilt.edu/math/graduate/](https://as.vanderbilt.edu/math/graduate/)

**Contact:** Denis Osin

**Description:** The most recent (2010) National Research Council (NRC) report on U.S. math graduate programs places our department in the top group of graduate programs surveyed. The Department of Mathematics has a distinguished international faculty that includes a Fields Medalist and International Congress of Mathematicians invited speakers. The department has a variety of research groups: universal algebra, topology, group theory, geometry/topology, approximation theory, noncommutative geometry, operator algebras, mathematical biology, partial differential equations, and graph theory. The stimulating research environment is supported by an ongoing program that attracts visiting scholars from institutions around the world and hosts several major conferences a year. The Department of Mathematics has approximately fifty research faculty and forty resident graduate students. It is large enough to support a wide range of courses, but small enough for students to receive individual attention from faculty members.

**Target Student:** The typical student has taken several upper-level math classes in their undergraduate study. Graduates of the PhD program are prepared for a wide range of career options, including academia and industry.

**Brief Curriculum Outline:** Doctoral candidates complete a core curriculum in algebra, topology, and analysis. After passing preliminary exams in two of these three areas, students study in their area of concentration, such as computational mathematics.

**Psychological Sciences, Quantitative Methods (A&S and Peabody), PhD**

**Web Site:** [https://www.vanderbilt.edu/psychological_sciences//graduate/programs/quantitative-methods/](https://www.vanderbilt.edu/psychological_sciences//graduate/programs/quantitative-methods/)

**Contact:** Daniel Levin (Peabody) / Geoff Woodman (A&S)

**Description:** Faculty in the Quantitative Methods (QM) program train students in state-of-the-art statistical methods and engage in research that develops and applies such methods. Students in the QM doctoral program develop expertise in the principles of research design and in the theoretical foundations and application of advanced statistical models for human behavior.

**Target Student:** The QM Ph.D. program prepares students for faculty positions in academic settings, methodology positions in basic or applied research centers, or methodology positions in industry.

**Brief Curriculum Outline:** Faculty in the Quantitative Methods (QM) program train students in state-of-the-art statistical methods and engage in research that develops and applies such methods. Students in the QM doctoral program develop expertise in the principles of research design and in the theoretical foundations and application of advanced statistical models for human behavior. Within the QM program, course offerings include item response theory (introductory and advanced), multilevel modeling, factor analysis, nonparametric statistics, multivariate statistics, structural equation modeling, latent class and mixture modeling, exploratory and graphical data analysis, latent growth curve modeling, psychological
measurement, statistical inference, correlation and regression, survival analysis, psychometric methods, individual differences, categorical data analysis, and experimental design.

Other Notes: Many students get an optional minor in biostatistics.

Doctor of Education in Leadership and Learning in Organizations (Peabody), Ed.D.
Web Site: https://peabodyonline.vanderbilt.edu/programs/edd-in-organizational-learning
Contact: Catherine Loss
Description: Online Ed.D. Across industries, effective leaders provide the support and direction necessary to foster development, leverage resources, create solutions and resolve complex systemic challenges. Vanderbilt’s online Ed.D. in leadership and learning in organizations develops results-oriented, forward-thinking professionals who can confidently drive systemic change.
Target Student: Mid-career professional with responsibility for organizational learning and change, across industries.
Brief Curriculum Outline: 15 credit hours of the Ed.D. is in Data and Analytics. Includes courses in Research Design, Strategy and Analytics, Applied Statistics, Data Science, Program Evaluation
Other Notes: This is an online Ed.D., through collaboration with 2U. Nearly all courses are entirely online, with the exception of a requirement for an annual on-campus immersion weekend.

Master’s Programs

Biomedical Informatics (Medicine), MS
Web Site: https://www.vumc.org/dbmi/research-ms-and-phd-program
Contact: Cynthia Gadd
Description: Vanderbilt DBMI’s MS and PhD in Biomedical Informatics is an internationally recognized research program, in which students actively engage in high quality research mentored by DBMI’s 60+ faculty. As a result, DBMI students have achieved outstanding publication records, as well as numerous recognitions from national organizations. Candidates with prior training in the computational and biosciences are preferred. Funding opportunities include our National Library of Medicine Biomedical Informatics Training Grant and postdoctoral research fellowships, as well as the VA Post Doctoral Fellowship in Medical Informatics & Quality Improvement.
Target Student: a student with interest in Biomedical Informatics
Brief Curriculum Outline: courses in bio and medical informatics foundations, machine learning, data analysis, and biology

Biostatistics (Medicine), MS
Web Site: https://www.vanderbilt.edu/biostatistics-graduate/current-program/ms-program/degree-requirements/
Contact: Jeffrey Blume
Description: Biostatisticians are interdisciplinary scientists who specialize in the interpretation of data as scientific evidence. They use mathematical frameworks to account for uncertainty, extract knowledge from data, and generalize results to target populations. The discipline of
Statistics laid the foundations for how science learns from data. This process, combined with recent advances in computing, has blossomed into the emerging field of “Data Science”. Vanderbilt’s Biostatistics program seeks to train the next generation of statistical leaders who are dedicated to the biomedical sciences and public health. They will design and conduct experiments, analyze data and interpret results, refine data capture techniques in mobile devices and images, and develop new statistical methods to support and encourage cutting edge scientific discoveries. Prediction algorithms developed by our students are already used for routine medical care here at Vanderbilt University and at other hospitals across the country. Many of these algorithms, and their novel underlying mathematical tools, will soon be available to you on your mobile device.

**Target Student:** student interested in statistics and clinical applications  
**Brief Curriculum Outline:** foundational and applied statistics courses

**Computer Science, MS**  
**Web Site:** [https://engineering.vanderbilt.edu/eecs/Graduate/Programs.php](https://engineering.vanderbilt.edu/eecs/Graduate/Programs.php)  
**Contact:** Akos Ledecki  
**Description:** The computer science program provides opportunities for students to join vibrant collaborative and interdisciplinary efforts in artificial intelligence, computer animation and virtual environments, cyber-physical systems, distributed real-time and embedded middleware, human-systems integration, image processing, intelligent learning environments, model-integrated computing, robotics, software engineering and trustworthy computing. Students can explore a broad range of research areas in computer science, as well as interdisciplinary research opportunities in areas related to the School of Engineering's strategic directions in health care and medicine, security, energy and natural resources and Data Science.  
**Target Student:** anyone with an aptitude for STEM.  
**Brief Curriculum Outline:** Core courses in computing, data structures, and algorithms, and elective/depth courses (AI, Databases, Big Data, Machine Learning) that allow a Data Science specialization

**Mathematics (A&S), MA**  
**Web Site:** [https://as.vanderbilt.edu/math/graduate/](https://as.vanderbilt.edu/math/graduate/)  
**Contact:** Denis Osin  
**Description:** The most recent (2010) National Research Council (NRC) report on U.S. math graduate programs places our department in the top group of graduate programs surveyed. The Department of Mathematics has a distinguished international faculty that includes a Fields Medalist and International Congress of Mathematicians invited speakers. The department has a variety of research groups: universal algebra, topology, group theory, geometry/topology, approximation theory, noncommutative geometry, operator algebras, mathematical biology, partial differential equations, and graph theory. The stimulating research environment is supported by an ongoing program that attracts visiting scholars from institutions around the world and hosts several major conferences a year. The Department of Mathematics has approximately fifty research faculty and forty resident graduate students. It is large enough to support a wide range of courses, but small enough for students to receive individual attention from faculty members.  
**Target Student:** This program is flexible and is particularly well suited for students preparing for careers in industry, actuarial work, or government.
Brief Curriculum Outline: Course Work – Thirty-six hours of graduate work, divided between major and minor subjects and approved by the Mathematics Department, must be taken. Up to twelve of these 36 hours may be in related fields such as computer science, economics, or physics.

Quantitative Methods (Peabody), M.Ed.
Web Site: https://peabody.vanderbilt.edu/departments/psych/med_quantitative_methods/index.php
Contact: Joseph Rodgers and Shane Hutton
Description: The master of education (M.Ed.) in Quantitative Methods (QM) is designed to provide students with strong quantitative methods training for applied research settings.
Target Student: Students for whom the new program would have interest and value are those who wish to work in school systems, government, industry, dedicated research institutes, academic settings, and medical school research settings.
Brief Curriculum Outline: Students in this 32-hour program take two required core courses in quantitative methods (PSY-GS 8861 & 8864, the two-semester introductory statistics sequence), two required hours of seminar activity, and eight additional courses, of which one may be a content course (i.e., outside the QM area) and one may be a QM course outside of Psychology and Human Development. Potential elective courses include: PSY-GS 8867, Multivariate Analysis; PSY-GS 8870, Correlation and Regression; PSY-GS 8873, Structural Equation Modeling; PSY-GS 8876, Psychological Measurement; PSY-GS 8850, Exploratory and Graphical Data Analysis; PSY-GS 8850, Nonparametric Statistics; PSY-GS 8879, Factor Analysis; PSY-GS 8882, Multilevel Modeling; PSY-GS 8888, Latent growth Curve Modeling; PSY-GS 8885, Latent Class and Mixture Modeling; PSY-GS 8850, Applied Bayesian Analysis for Latent Variable Modeling; PSY-GS 8880, Introduction to Item Response Theory; PSY-GS 8881, Advanced Item Response Theory. The program culminates in a summer-long or semester-long internship in which students obtain real-world experience producing data analyses for a public or private organization in Nashville or the broader research community.

College of Arts and Science Humanities and Divinity
Notes: Not a degree program, but many humanities graduate students interested in Digital Humanities add working with digital humanities data as a competence through training in VU’s HASTAC Scholars program (HASTAC = Humanities, Arts, Science, and Technology Alliance and Collaboratory). See https://www.vanderbilt.edu/digitalhumanities/hastac-2017-2018/ On an informal level, many graduate students in the humanities train on using digital data through the ongoing digital scholarship workshops and working groups run by Jean and Alexander Heard Library staff. See http://library.vanderbilt.edu/scholarly/working-groups.php (for GIS, Linked Data and the Semantic Web, Tiny Data, and XQuery.) and http://library.vanderbilt.edu/scholarly/workshops.php

4+1 BA/MA and BS/MS Programs

College of Arts and Science 4+1 Program, BA/MA
https://as.vanderbilt.edu/academics/specialdegree/4plus1.php
Computer Science, BS/MS
https://engineering.vanderbilt.edu/eecs/Undergraduate/ (look for 5 year BS/MS program; it doesn’t have an independent website)

Undergraduate Programs

Computer Science, Major/Minor (Engineering or A&S)
Web Site: https://my.vanderbilt.edu/juliejohnson/dus_info/
Contact: Julie Johnson
Description: If you major in computer science or computer engineering, you will have a broad range of career choices. Computer scientists design computer software and use computer programming to solve a wide range of problems. Computer engineers organize, design, and apply digital processing systems, bridging hardware and software.
Target Student: anyone with an aptitude for STEM.
Brief Curriculum Outline: Core courses in computing, data structures, and algorithms, and elective/depth courses (AI, Databases, Big Data, Projects in AI, Machine Learning) that allow a Data Science specialization

Mathematics (A&S), Major/Minor
Web Site: https://as.vanderbilt.edu/math/graduate/
Contact: John Rafter
Description: The Department of Mathematics offers an undergraduate major with a high degree of flexibility. A solid background in mathematics provides an excellent foundation for any quantitative discipline as well as many professions—many students go on to professional studies in law, medicine, or business.
Brief Curriculum Outline: Three tracks: Standard, Applied, Honors, require sequences in calculus, linear algebra, and differential equations, as well as several upper level courses, which could be statistics courses.

Quantitative Social Sciences (A&S and Peabody), Major
Web Site: (proposed program)
Contact: Alan Wiseman (Political Science)
Description: The Program in Quantitative Social Sciences (QSS) is an interdisciplinary program that provides students with a strong foundation in mathematics, the tools of quantitative analysis, and a substantive area of expertise in a social science discipline. The Program offers a major. Through course offerings, colloquia, and research opportunities, program faculty help students develop their analytic skills in multivariate calculus, linear algebra, probability theory, regression, econometrics, game theory, network analysis, and formal modeling. Students choose a discipline-specific area of concentration, developing discipline-specific knowledge and perspectives.
Target Student: Aims to facilitate the creation of a new cohort of Vanderbilt undergraduate students who, upon graduation, would be equipped with a strong foundation in mathematics and the tools of quantitative analysis in addition to a substantive area of expertise, typically in economics, political science, psychology, or sociology.
Brief Curriculum Outline: Core courses: QSS 1001 Introduction to Scientific Methods or QSS 1111-1 Big Data, Scientific Methods, and Quantitative Analyses; QSS 1001L Introduction to Scientific Methods Computing Lab; Math 2310 Multivariable Calculus & Linear Algebra; One of the following: Math 2310, Math 2300 and Math 2600, Math 2300 and Math 2410, Math 2500 and Math 2501, Math 2820 Introduction to Probability Theory and Mathematical Statistics; QSS 2820L Probability Theory and Mathematical Statistics for QSS Computing Lab; QSS 3001 Regression Analysis for QSS or PSY-PC 3735, Correlation and Regression. Wide range of electives from HOD, PSY, ECON, and the like. Plus a disciplinary focus in ECON, PSCI, PSY, PSY-PC, SOC.

Scientific Computing (A&S and Engineering), Minor
Web Site: https://www.vanderbilt.edu/scientific_computing/
Contact: Thomas Palmeri, Robert Bodenheimer, or David Weintraub
Description: Students in the program in Scientific Computing are taught techniques for understanding complex physical, biological, and social systems. Students are introduced to computational methods for simulating and analyzing models of complex systems, to scientific visualization and data mining techniques needed to detect structure in massively large multidimensional data sets, to high performance computing techniques for simulating models on computing clusters with hundreds or thousands of parallel, independent processors and for analyzing terabytes or more of data that may be distributed across a massive cloud or grid storage environment.
Target Student: Student in social sciences, sciences, or engineering who want to do computational work
Brief Curriculum Outline: Intro programming (CS 1101), data structures (CS 2204), Scientific Computing Toolbox (SC 3250), broad selection of electives across 17 departments in A&S, Engineering, and the Medical Center
APPENDIX C: Vanderbilt Courses Related to Data Science

**Introductory Programming**
- CS 1000 Introduction to Computation
- CS 1101 Programming and Problem Solving
- CS 1103 Introductory Programming for Engineers and Scientists
- CS 2201 Program Design and Data Structures
- CS 2204 Program Design and Data Structures for Scientific Computing
- PSY 8219 Scientific Computing for Psychological and Brain Science

**Advanced Programming**
- CS 3270/4270 Programming Languages
- CS 4250/5250 Algorithms
- CS 6310 Design and Analysis of Algorithms
- CS 6320 Algorithms for Parallel Computing
- SC 3260/5260 High Performance Computing
- SC 3890/5890 Special Topics in Scientific Computing: Advanced High Performance Computing

**Introductory Statistics**
- BIOS 6301 Introduction to Statistical Computing
- BIOS 6311 Principles of Modern Biostatistics
- BSCI 5270 Statistical Methods in Biology
- ECON 1500 Economic Statistics
- PSY 2100 Quantitative Methods
- PSY-PC 2110 Introduction to Statistical Analysis
- PSY-PC 2120 Statistical Analysis
- MATH 1010/1011 Probability and Statistical Inference
- MATH 2810 Probability and Statistics for Engineering
- MATH 2820/5820 Introduction to Probability and Mathematical Statistics
- MATH 2820L Statistics Laboratory
- MATH 2821/5821 Introduction to Applied Statistics
- QSS 1001 Introduction to Scientific Methods
- QSS 1111 Big Data, Scientific Methods, and Quantitative Analyses
- QSS 1001L Introduction to Scientific Methods Computing Lab

**Advanced Statistics**
- BIOS 6312 Modern Regression Analysis
- BIOS 6321 Clinical Trials and Experimental Design
- BIOS 6341 Fundamentals of Probability
- BIOS 6342 Contemporary Statistical Inference
- BIOS 7323 Applied Survival Analysis
- BIOS 7330 Regression Modeling Strategies
- BIOS 7345 Advanced Regression Analysis I (Linear & General Linear Models)
BIOS 7346 Advanced Regression Analysis II (General Linear Models & Longitudinal Data Analysis)
BIOS 7361 Advanced Probability and Real Analysis Concepts
BIOS 7362 Advanced Statistical Inference and Statistical Learning
BIOS 8366 Advanced Statistical Computing
BIOS 8370 Foundations of Statistical Inference
BIOS 8372 Bayesian Methods
BIOS 8375 Causal Inference
CE 6310 Uncertainty Quantification
CE 3300 Risk, Reliability, and Resilience Engineering
CE 3890/5999 Data Analytics for Engineers
CE 4300/5300 Reliability and Risk Case Studies
ECON 3032 Applied Econometrics
ECON 3035 Econometric Methods
ECON 3050 Introduction to Econometrics
ECON 3330 Economics of Risk
ECON 4050/5050 Topics in Econometrics
ECON 6600 Econometrics
ECON 8300 Statistical Analysis
ECON 8310 Econometrics I
ECON 8320 Econometrics II
ECON 9310 Time Series Econometrics
ECON 9320 Non-parametric and Semi-parametric Econometrics
ECON 9330 Microeconometrics
ECON 9710 Public Economics: Expenditure
ECON 9720 Public Economics: Taxation
EECE 5287 Engineering Reliability
HOD 8850/3275 Practical Meta-analysis
HOD 3205 Education Policy Analysis Methods
LPO 7810 Causal Inference
MATH 3640/5640 Probability
MATH 3641/5641 Mathematical Statistics
MATH 4650 Financial Stochastic Processes
MATH 4651 Evaluation of Actuarial Models
PSCI 8356 Statistics for Political Research I
PSCI 8357 Statistics for Political Research II
PSCI 8362 Data Collection Methods
PSCI 8363 Survey Research Methods
PSY 3891/8551 Special Topics in Cognitive Psychology - Bayesian Cognitive Modeling
PSY 6104 Quantitative Methods and Experimental Design
PSY 8120 Categorical Data Analysis
PSY 8864 Experimental Design
PSY 8858 Introduction to Statistical Inference
PSY 8305 Linear and Nonlinear Mixed Effects Models
PSY-GS 3738/8880 Introduction to Item Response Theory
PSY-GS 8881 Item Response Theory II
PSY-GS 8882 Multilevel Modeling
PSY-GS 3743/8879 Factor Analysis
PSY-GS 3749/8850 Applied Nonparametric Statistics
PSY-GS 3746/8867 Multivariate Statistics
PSY-GS 8850 Advanced Structural Equation Modeling
PSY-GS 3730/8885 Applied Latent Class and Mixture Modeling
PSY-GS 3737/8873 Structural Equation Modeling
PSY-PC 3732/8888 Latent Growth Curve Modeling
PSY-GS 3724/8876 Psychological Measurement
PSY-GS 8861 Statistical Inference
PSY-GS 3735/8870 Correlation and Regression
PSY-GS 8850 Survival Analysis
PSY-GS 8855 Quantitative Methods Forum
PSY-PC 3727 Modern Robust Statistical Methods
PSY-PC 3722 Psychometric Methods
SOC 6311 Multivariate Analysis I
SOC 6312 Multivariate Analysis II
SOC 7600 Quantitative Methods Workshop

Mathematical Fundamentals
MATH 2410 Methods of Linear Algebra
MATH 2600/5600 Linear Algebra
MATH 3130 Fourier Analysis
MATH 3620 Introduction to Numerical Mathematics
MATH 3890 Advanced Linear Algebra
MATH 4600 Numerical Analysis
MATH 4620 Linear Optimization
MATH 4630 Nonlinear Optimization
MATH 4710/6710 Graph Theory
MATH 7130 Harmonic Analysis
MATH 9600 Computing with Splines

Visualization
ANTH 3261 Introduction to Geographic Information Systems and Remote Sensing
EECE 4353/5353 Image Processing
EECE 6358 Quantitative Medical Image Analysis
PSY-GS 3751/8850 Exploratory and Graphical Data Analysis
EE 3892/5892 Special Topics course on Statistical Pattern Recognition

Data Science
HOD 3200 Introduction to Data Science
MATH 3670/5670 Mathematical Data Science
PSCI 3893 Data Science for Politics

Big Data
CS 4266/4266 Topics in Big Data
CS 4287/5287 Principles of Cloud Computing

**Informatics / Bioinformatics / Databases**
BMIF 6300 Foundations of Biomedical Informatics
BMIF 6310 Foundations of Bioinformatics
BMIF 7320 Healthcare System and Informatics
CS 3265/5265 Introduction to Database Management Systems
BMIF 7340 Clinical Information Systems and Databases

**Data Privacy / Ethics**
BMIF 7380 Data Privacy in Biomedicine
CS 1151 Computers and Ethics

**Machine Learning**
BMIF 7330 Machine Learning for Biomedicine
CS 3891/5891 Deep Learning
CS 4260/5260 Artificial Intelligence
CS 6362 Machine Learning
CS 6350 Artificial Neural Networks
CS 6360 Advanced AI
CS 6364 Intelligent Learning Environments
EECE 4354/5354 Computer Vision
NSC 3270 Computational Neuroscience

**Computation / Simulation**
ASTR 3600/8030 Stellar Astrophysics
ASTR 3700 Galactic Astrophysics
ASTR 3800/8050 Large Scale Structure in the Universe
BSCI 3272 Genome Science
EECE 3214 Signals and Systems
EECE 4252/5252 Signal Processing and Communications
EECE 4356 /5356 Digital Signal Processing
EECE 6361 Random Processes
EECE 6362 Detection and Estimation Theory
MATH 3630 Mathematical Modeling in Biology
MATH 3660 Mathematical Modeling in Economics
PHYS 3200/5200 Thermal and Statistical Physics
PHYS 3645/7645 Radiation Detectors and Measurements
PHYS 3790/5237 Computational Physics
PHYS 8105 Special Topics in Experimental Physics
PHYS 8124 Physical Measurements on Biological Systems
PHYS 8126 Theoretical and Experimental Systems Biology
PSY 8218 Computational Modeling
PSY 8503 Models of Human Memory
SC 3250/5250 Scientific Computing Toolbox
This appendix lists the results of a survey of data science expertise among Vanderbilt faculty. Faculty were asked to self-identify as expert (E) or frequent user (U) of a core set of data science methodologies: machine learning, data visualization, data engineering, statistics, data simulation, or policy and ethics relating to data. In this context, “expert” means sufficiently knowledgeable to teach a course on the topic. Faculty were also asked to self-identify as experienced users (U) of different types of data: natural language, images, video, audio, geospatial, time series, relational database, or clickstream. Several faculty also listed methods or types of data that were not included in the original survey, such as “mathematical modeling” or “genome data”.

Responses were collected from over 300 faculty representing 41 academic departments in all eight of Vanderbilt’s schools and colleges. The survey is reasonably complete, though there are a few departments missing, most of which are clinical departments in VUMC. There are also undoubtedly some faculty missing from departments that are listed and some listed faculty whose information is inaccurate. This survey should thus be treated as a work in progress.
<table>
<thead>
<tr>
<th>Name</th>
<th>Machine Learning</th>
<th>Data Visualization</th>
<th>Data Engineering</th>
<th>Statistics</th>
<th>Data Simulation &amp; Ethics</th>
<th>Natural Language</th>
<th>Images</th>
<th>Video</th>
<th>Audio</th>
<th>Geo-spatial</th>
<th>Time Series</th>
<th>Relational Database</th>
<th>Click-stream</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anthropology</strong></td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benn-Torres</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ross</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tung</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Werneke</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Biological Sciences</strong></td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baskauf</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bordensain</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brodie</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capra</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creanza</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hilburn</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McMahon</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rokas</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tate</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Chemistry</strong></td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachman</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McLean</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miller</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lybrand</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Classical and Mediterranean Studies</strong></td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rife</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Earth and Environmental Sciences</strong></td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ayers</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bennartz</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Damoch</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Furbish</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gilligan</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kelly</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Economics</strong></td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carpenter</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collins</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crucini</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dustan</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goodman-Bacon</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gutierez</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inoue</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marcus</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moro</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rodrigue</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rousseau</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SantAnna</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sasaki</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zargoza-Watkins</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zimnan</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>English</strong></td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meadows</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clayton</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shin</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>French</strong></td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramsey</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>History</strong></td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moleviv</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landers</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Igo</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>History of Art</strong></td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miller</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mathematics</strong></td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aldroubi</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ellingham</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hamm</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hartin</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neamtu</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Powell</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raffi</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saff</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schumaker</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Webb</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Machine Learning</td>
<td>Data Visualization</td>
<td>Data Engineering</td>
<td>Statistics</td>
<td>Data Simulation</td>
<td>Policy &amp; Ethics</td>
<td>Natural Language</td>
<td>Images</td>
<td>Video</td>
<td>Audio</td>
<td>Geo-spatial</td>
<td>Time Series</td>
<td>Relational Database</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------</td>
<td>--------------------</td>
<td>------------------</td>
<td>------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------------------</td>
</tr>
<tr>
<td><strong>Physics and Astronomy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berland</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greene</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guerola</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holley-Bockelmann</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holmes</td>
<td></td>
<td>E</td>
<td>E</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hutson</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johns</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rericha</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheehan</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stassun</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Velkovska</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wikswo</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Political Science</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clayton</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinton</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kam</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kohnke</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Larsen</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lupu</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michelitch</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schram</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smith</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zachmeister</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Psychology</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachorowski</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gauthier</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hoffman</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logan</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maier</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marcis</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McNamara</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palmeri</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polyn</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schall</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tomarkien</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tong</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trueblood</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td></td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Womelsdorf</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zaid</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Religious Studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sociology</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carpenter</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christie-Mizell</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cornfield</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erving</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hess</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hess</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isaac</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keler</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lloyd</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McComman</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murray</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patterson</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitt</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sana</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Song</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trotter</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SCHOOL OF ENGINEERING</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Biomedical Engineering</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Byron</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Civil and Environmental Engineering</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Askewitz</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adams</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baroud</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hornberger</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krahm</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lin</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Electrical Engineering and Computer Science</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

51
<table>
<thead>
<tr>
<th>Name</th>
<th>Machine Learning</th>
<th>Data Visualization</th>
<th>Data Engineering</th>
<th>Data Statistics</th>
<th>Data Simulation</th>
<th>Policy &amp; Ethics</th>
<th>Natural Language</th>
<th>Images</th>
<th>Video</th>
<th>Audio</th>
<th>Geo-spatial</th>
<th>Time Series</th>
<th>Relational Database</th>
<th>Click-stream</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biswas</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dawant</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dubey</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fisher</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koutoukos</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kunda</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landman</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td>E</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noble</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peters</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rodriguez-Cancio</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schmidt</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vorobeychik</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td></td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mechanical Engineering</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luo</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valentine</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walker</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SCHOOL OF MEDICINE AND VUMC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Anesthesiology</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wanderer</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Biochemistry</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blind</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hodges</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heibert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chazin</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ren</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bowman</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lacy</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ren</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lopez</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quaranta</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Biomedical Informatics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bejan</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cao</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carroll</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chan</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cronin</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denny</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duda</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fabini</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farishek</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cobbol</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giuse</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harris</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hughey</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lasko</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Levy</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liu</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malin</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matheny</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paterson</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reeves</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosenthal</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smith</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sternier</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walsh</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wei</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yin</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td></td>
<td>E</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bioinformatics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blume</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Byrne</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choi</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fonnesbeck</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hackstadt</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td></td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harrell</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kang</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koyama</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liu</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Machine</td>
<td>Data Visualization</td>
<td>Data Engineering</td>
<td>Data Statistics</td>
<td>Data Simulation &amp; Policy</td>
<td>Data Ethics</td>
<td>Natural Language</td>
<td>Images</td>
<td>Video</td>
<td>Audio</td>
<td>Geo-spatial</td>
<td>Time Series</td>
<td>Relational Database</td>
<td>Click-stream</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------</td>
<td>--------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>--------------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>--------------</td>
<td>---------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Samuels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schildcrout</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheng</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shepherd</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shotwell</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smith</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stewart</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tao</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell and Developmental Biology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burnette</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gould</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irlie</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irish</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kavolina</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lau</td>
<td>E</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macara</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miller</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page-McCaw</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tanser</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyska</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weaver</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zamic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McNaughton</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Policy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beskow</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dusetzina</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Everson</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gonzales</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graves</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Griffin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hamad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koeheke</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nikpay</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patrick</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ray</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schaffner</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiese</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cox</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Davis</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Velez-Edwards</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edwards</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warner</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Molecular Physiology and Biophysics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuert</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samuels</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sutcliffe</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pathology, Microbiology and Immunology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Georgiev</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgical Sciences (8 departments)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aldrich</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bichard</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depen</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jackson</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pousoae</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peabody College of Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human and Organizational Development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barnes</td>
<td>E</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christians</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cobb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diah</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erford</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forber-Pratt</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraser</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friedman</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marin</td>
<td>U</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Machine Learning</td>
<td>Data Visualization</td>
<td>Data Engineering</td>
<td>Statistics</td>
<td>Data Simulation</td>
<td>Policy &amp; Ethics</td>
<td>Natural Language</td>
<td>Images</td>
<td>Video</td>
<td>Audio</td>
<td>Geo-spatial</td>
<td>Time Series</td>
<td>Relational Database</td>
<td>Click-stream</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------</td>
<td>--------------------</td>
<td>------------------</td>
<td>------------</td>
<td>----------------</td>
<td>----------------</td>
<td>------------------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>McBride</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>McDonald</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Nation</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Nixon</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Nolan</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Perkins, D</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Perkin, J</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Shinn</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Speer</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Suiter</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Van Schaack</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td><strong>Leadership, Policy, and Organizations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ballou</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candestalat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doyle</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Evans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grisoom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heinrich</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harvey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Psychology and Human Development</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cho</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hutton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preacher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rodgers</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sterba</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Special Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yoder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Teaching and Learning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brady</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Durkin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leander</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PROFESSIONAL SCHOOLS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Blair School of Music</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calico</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barz</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Divinity School</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michelson</td>
<td>E</td>
<td>U</td>
<td>E</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Byrd</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hammon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armour</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheppard</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Townes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Law School</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cheng</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Ruhl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>George</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slabogin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>School of Nursing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piano</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dietrich</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phillipi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mulvany</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weiner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ahard</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maxwell</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhoden</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rider</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Owen School of Management</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goldsmith</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stuart</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chaney</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cool</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whaley</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bollen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blocher</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Munyan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lewis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensoy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Machine Learning</td>
<td>Data Visualization</td>
<td>Data Engineering</td>
<td>Statistics</td>
<td>Data Simulation &amp; Analytics</td>
<td>Policy &amp; Ethics</td>
<td>Natural Language</td>
<td>Images</td>
<td>Video</td>
<td>Audio</td>
<td>Geo-spatial</td>
<td>Time Series</td>
<td>Relational Database</td>
<td>Clickstream</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------</td>
<td>--------------------</td>
<td>------------------</td>
<td>------------</td>
<td>-----------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>--------</td>
<td>-------</td>
<td>------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Haslag</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parsley</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hoefler</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iacobucci</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Willis</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posavec</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christie</td>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LeBlanc</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thur</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Experts</td>
<td>33  22  18  84  56  14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Experienced Users</td>
<td>58  96  61  114  44  47</td>
<td>56  72  33  26  48  93  76  39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX E: Data Science Programs at Peer Institutions

This appendix lists education programs in data science at the top 30 research universities (plus a few peer institutions in the top 70) and top 20 liberal arts colleges (according to the US News & World Report 2018 rankings). The following table shows the school name and rank (column 1), type of statistics department present (column 2), whether there is a biostatistics department (column 3), and types of degrees or certifications available in data science (columns 4–9).

<table>
<thead>
<tr>
<th>Research Universities</th>
<th>Statistics (S) or Mathematics &amp; Statistics (M)</th>
<th>Bio-statistics</th>
<th>PhD</th>
<th>MS</th>
<th>Grad Cert</th>
<th>BA/BS</th>
<th>U Minor</th>
<th>U Cert</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Princeton</td>
<td>Applied Mathematics (A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Harvard</td>
<td>Statistics (S)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 UChicago</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Yale</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Columbia</td>
<td>S</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 MIT</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Stanford</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 UPenn</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Duke</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Cal Tech</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Dartmouth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Johns Hopkins</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Northwestern</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Brown</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Cornell</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Rice</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Vanderbilt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Notre Dame</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 WashU in St Louis</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 Georgetown</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 Emory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 UC Berkeley</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 UCLA</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 U of Southern Cal</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 CMU</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 Virginia</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27 Wake Forest</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28 UMichigan</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29 Tufts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 NYU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 UNC, Chapel Hill</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34 URochester</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34 Georgia Tech</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46 Wisconsin</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54 Ohio State</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56 Univ of Washington</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69 Minnesota</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69 Rutgers</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liberal Arts Colleges</td>
<td>Statistics (S) or Mathematics &amp; Statistics (M)</td>
<td>Bio-statistics</td>
<td>PhD DS</td>
<td>MS DS</td>
<td>Grad Cert DS</td>
<td>BA/BS DS</td>
<td>U Minor DS</td>
<td>U Cert DS</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>--------</td>
<td>-------</td>
<td>-------------</td>
<td>---------</td>
<td>-----------</td>
<td>---------</td>
</tr>
<tr>
<td>1 Williams</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Amherst</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Bowdoin</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Swarthmore</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Wellesley</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Middlebury</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Pomona</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Carleton</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>8 Claremont McKenna</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Davidson College</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>12 Colby</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>12 Colgate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 Harvey Mudd</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 Smith College</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>12 US Military Academy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 Vassar</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Grinnell</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Hamilton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Haverford</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
This appendix provides more in-depth information on educational programs in data science for a subset of peer institutions: Columbia University, University of Pennsylvania, University of Chicago, University of Virginia, Brown University, New York University, University of Rochester, and University of Washington. For each program, there are answers to twelve questions that address the basic structure and nature of the program.

**Columbia University**

**[Columbia] Undergraduate (BA in Data Science)**

BA in Data Science as a part of the CS department. Notably, this is a BA, not a BS, and combines computer science and statistics core and elective requirements. Heavily focused on foundations, rather than applications.

1. **is it a standalone program or a track within another department/program - is it intended to be connected with another major/program**
   - track within CS, but strongly connected with statistics
2. **does the program have a specific focus (e.g., business analytics, bio data) or more general**
   - general
3. **does the program have multiple tracks, emphasis areas, and if so what - are there different threads**
   - no, it is a single-track program within CS
4. **do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)**
   - no
5. **does the program have more of a computational, mathematical, statistical emphasis**
   - heavy computational and statistical emphasis
6. **what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)**
   - the program relies on coursework in mathematical foundations for CS and statistics – it is a CS program, so standard programming foundations for CS are required.
7. **for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect**
   - N/A
8. **what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)**
   - applied Data Science, applied Data Mining, AI/ML courses
9. **how centralized is the center/program in which the program resides**
   - centralized (in CS), but coursework across computer science and statistics
10. **are there capstone projects**
    - no
(11) are there internships (maybe along with capstone projects)
no
(12) how applied? how theoretical?
quite theoretical/foundations-based

[Columbia] Masters (MS in Data Science)
http://datascience.columbia.edu/data-science-academics
Columbia University Data Science Institute offers interdisciplinary data science graduate programs with part-time, full-time and online study options.

(1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
connected with Columbia University Data Science Institute
(2) does the program have a specific focus (e.g., business analytics, bio data) or more general
general
(3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
single track
(4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
no – the program focus is general data science
(5) does the program have more of a computational, mathematical, statistical emphasis
yes, focus is on computing and statistics
(6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
Requires an undergraduate degree with prior quantitative coursework (i.e., linear algebra, probability/statistics, etc.); prior introductory to computer programming coursework (i.e., Python, Java, C++, etc.). Candidates for the Master of Science in Data Science are required to complete a minimum of 30 credits, including 21 credits of required/core courses and 9 credits of electives. This program may be pursued part-time or full-time. Most students enroll on a full-time status, completing the program in three semesters/one and a half years (Fall: 12-credits; Spring: 12-credits; Summer: Optional Internship or elective; Fall: final 3- or 6-credits). Core courses in Statistics and CS. Electives more open, but subject to approval (http://datascience.columbia.edu/course-inventory).
(7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
expects a relatively rigorous technical background in math, statistics, and computing
(8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
Introduction to Data Science, Computer Systems for Data Science, Algorithms for Data Science, Machine Learning for Data Science
(9) how centralized is the center/program in which the program resides
the program appears relatively centralized in the Data Science Institute, but coursework across computer science and statistics
(10) are there capstone projects
Data Science capstone and ethics course (required)
(11) are there internships (maybe along with capstone projects)
optional internship
(12) how applied? how theoretical?
focus is on foundations rather than applications, but the program has more applied elective courses, such as Natural Language Processing and Bioinformatics

[Columbia] Certificate in Data Science
http://datascience.columbia.edu/data-science-academics
The Certification of Professional Achievement in Data Sciences prepares students to expand their career prospects or change career paths by developing foundational data science skills. Join from anywhere in the world as the program is now also offered online as of the Fall 2016 term.

(1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
connected with Columbia University Data Science Institute
(2) does the program have a specific focus (e.g., business analytics, bio data) or more general
general
(3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
single track
(4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
no, but assumes technical background (math, linear algebra, programming)
(5) does the program have more of a computational, mathematical, statistical emphasis
yes, focus on computing and statistics
(6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
Candidates for the Certification of Professional Achievement in Data Sciences, a non-degree part-time program, are required to complete a minimum of 12 credits, including four required courses: in statistics, algorithms, systems, and machine learning. Broader elective options (http://datascience.columbia.edu/course-inventory).
(7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
undergraduates with math, statistics, and programming background
(8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
N/A
(9) how centralized is the center/program in which the program resides
the program appears relatively centralized in the Data Science Institute, but coursework across computer science and statistics
(10) are there capstone projects
no
(11) are there internships (maybe along with capstone projects)
   no
(12) how applied? how theoretical?
   relatively theoretical

University of Pennsylvania

[U Penn] Undergraduate (Data Science minor)
https://catalog.upenn.edu/undergraduate/programs/data-science-minor/
Data Science applies core concepts in computer science, statistics and mathematics to problems in a wide variety of fields, from physical, social, biomedical, and behavioral sciences to arts and humanities. The minor targets students with strong analytical abilities and some existing programming experience, and requires courses in statistics, data-centric programming, data management, and data analysis. It also points to courses across the University that deal with data in areas of importance to Data Science.

(1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
   a minor, connected with any other major
(2) does the program have a specific focus (e.g., business analytics, bio data) or more general general
(3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
   no, it is a single-track program
(4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
   no
(5) does the program have more of a computational, mathematical, statistical emphasis
   heavy computational and statistical emphasis
(6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
   the curriculum includes relevant math and programming prerequisites. CIS 120 Programming Languages and Techniques I (http://www.seas.upenn.edu/~cis120/current/); students with no prior programming experience are recommended (but not required) to take CIS 110 (similar to our CS 1101), CIS 419 Introduction to Machine Learning OR STAT 471 Modern Data Mining, selection of UG statistics courses, such as ENM 321 Engineering Statistics; this one has no prerequisites
(7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
   N/A
(8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
largely relies on existing courses; Introduction to Large-Scale Data Science is perhaps the only special course

(9) how centralized is the center/program in which the program resides
relatively decentralized, housed in the School of Engineering, but cross-disciplinary with stats, math, marketing, operations, broader engineering

(10) are there capstone projects
no

(11) are there internships (maybe along with capstone projects)
no

(12) how applied? how theoretical?
quite theoretical/foundations-based

[UPenn] Undergraduate (Business Analytics major in the Wharton School of Business)
https://statistics.wharton.upenn.edu/programs/undergraduate/business-analytics-joint-concentration/
The Business Analytics joint concentration between the OID and STAT departments is designed to build deep competency in the skills needed to implement and oversee data-driven business decisions, including (i) collecting, managing and describing datasets, (ii) forming inferences and predictions from data and (iii) making optimal and robust decisions. Business analytics makes extensive use of statistical analysis and the applications of business analytics span all functional areas.

(1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
standalone program in Wharton Business School

(2) does the program have a specific focus (e.g., business analytics, bio data) or more general specific: Business Analytics

(3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
no, it is a single-track program

(4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
no

(5) does the program have more of a computational, mathematical, statistical emphasis
heavy computational and statistical emphasis

(6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)

Beyond core course requirements in OIDD and STAT, students must complete 4 credit units from the set of approved courses listed below. Furthermore, among the set of selected courses, there must be at least one course that provides competency for each of the following three fundamental skills in business analytics: Data collection: methods for acquiring and manipulating data, Advanced data analysis: working with data sets in a computing environment, Optimization: computer-based prescriptive decision making

(7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
(8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
    largely relies on existing courses
(9) how centralized is the center/program in which the program resides
    centralized
(10) are there capstone projects
    no
(11) are there internships (maybe along with capstone projects)
    no
(12) how applied? how theoretical?
    quite theoretical/Foundations-based

[UPenn] Master’s (MS in Engineering and Data Science)
https://dats.seas.upenn.edu
The Data Science degree program can be completed in one to two years. It blends leading-edge courses in core topics such as machine learning, big data analytics, and statistics, with a variety of electives and an opportunity to apply these techniques in a domain specialization – a depth area – of choice. The depth area offers both preparatory coursework and a thesis or practicum in a data science application area. Potential areas of specialization include network science (the Warren Center for Network and Data Science), digital humanities (the Price Lab for Digital Humanities), biomedicine (the Institute for Biomedical Informatics), and public policy (the Annenberg Center for Public Policy) — as well as more traditional opportunities in Computer and Information Science and Electrical and Systems Engineering. For students interested in applying data analysis and modeling to other areas within engineering and the physical sciences, Penn offers a specialized and synergistic program in Scientific Computing.

(1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
    standalone program in the School of Engineering
(2) does the program have a specific focus (e.g., business analytics, bio data) or more general
    general
(3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
    no, it is a single-track program
(4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
    no
(5) does the program have more of a computational, mathematical, statistical emphasis
    heavy computational and statistical emphasis, but electives can cover a broad array of courses
(6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
    The 10 Course Units in the degree are divided into four categories. Foundations (2 courses): Graduate probability and statistics (from Engineering, Stat, or Math), Programming
languages / intro to programming from Upenn CS (technically, CIS/CIT) (can be waived with requisite background)

Core (3 courses): Stats/Optimization: Linear algebra/optimization, Computational Learning Theory, or Mathematical Statistics; Machine Learning; Big Data Analytics

Electives: 3 courses; 1 approved technical depth, and two from a Depth Area in CS, Systems Engineering, Network/Social Science, Digital Humanities, Biomedicine (with approval of program director); Thesis/practicum: an approved 2-course sequence, or a year-long thesis project

(7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect

undergraduates with or without prior math/stats/CS background

(8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)

largely relies on existing courses

(9) how centralized is the center/program in which the program resides

centralized (School of Engineering), but combines classes from CS and Statistics

(10) are there capstone projects

no, but there is a thesis option

(11) are there internships (maybe along with capstone projects)

no

(12) how applied? how theoretical?

quite theoretical/ foundations-based, but electives can include more applied courses, such as in digital humanities and biomedicine

[UPenn] Master’s (MBA in Business Analytics)
https://oid.wharton.upenn.edu/programs/mba/business-analytics/

The Business Analytics MBA major is designed to build deep competency in the skills needed to implement and oversee data-driven business decisions, including (i) collecting, managing and describing datasets, (ii) forming inferences and predictions from data, and (iii) making optimal and robust decisions. Business analytics makes extensive use of statistical analysis, and the applications of business analytics span all functional areas.

(1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program

standalone program in Wharton Business School

(2) does the program have a specific focus (e.g., business analytics, bio data) or more general specific: Business Analytics

(3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads

no, it is a single-track program

(4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)

no

(5) does the program have more of a computational, mathematical, statistical emphasis
heavy computational and statistical emphasis, but electives can cover a broad array of courses

(6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)

In addition to OIDD 611 or OIDD 615, students must complete OIDD 612 as part of their OIDD flexible core requirement. Students must also complete the required fixed core requirement in STAT. Waiving the STAT fixed core requirement entirely or placing into and passing STAT 621 fulfills the STAT fixed core requirement for this major. In addition, students must complete 4 credit units from the set of approved courses

(7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect

MBA students with an interest in business analytics. Does not appear to assume technical background, but coursework is quite technical

(8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)

largely relies on existing courses

(9) how centralized is the center/program in which the program resides

centralized (Wharton), but broad selection of coursework across the Business School (includes OIDD, Marketing, and Statistics)

(10) are there capstone projects

no

(11) are there internships (maybe along with capstone projects)

no

(12) how applied? how theoretical?

focused on business applications, rather than data science foundations

[UPenn] Doctoral (PhD in Applied Mathematics and Computational Science)

https://www.amcs.upenn.edu/academic-programs/phd-program

The degree of Doctor of Philosophy in Applied Mathematics and Computational Science is an advanced degree designed for those who wish to pursue a career involving applied mathematics research. It is conferred in recognition of marked ability and high attainment in advanced applied and computational mathematics, including the successful completion of a significant original research project. The Ph.D. program is designed to guide students, year-by-year, toward becoming researchers in applied mathematicians. Typically the program takes four or five years to complete, including the dissertation (although it can be completed in less time, depending on the student).

(1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program

standalone program

(2) does the program have a specific focus (e.g., business analytics, bio data) or more general

general

(3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
no, it is a single-track program
(4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
no
(5) does the program have more of a computational, mathematical, statistical emphasis
yes
(6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
Twenty units of graduate courses required for the Ph.D. degree. The bulk of these courses should be drawn from the list of approved courses. (Independent study courses at Penn may be counted toward the twenty course requirement.) Among the courses, every student must take at least two semesters of graduate courses at the 600 level in each of applied algebra, and applied analysis, at least one semester of probability and stochastic processes, and one semester of computational science. In general, eight of the courses should be taken in AMCS itself or in the Mathematics department. After passing the oral exam/thesis proposal requirement, PhD students should register for 3CUs of dissertation research (AMCS 999) until they have accumulated the required 20 CUs. Once they have fulfilled this require they can register for AMCS 995, which carries no credits. To receive the Ph.D. degree a student must have at least a 3.0 cumulative grade point average.
(7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
students with a strong math and CS background
(8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
largely relies on existing courses
(9) how centralized is the center/program in which the program resides
centralized
(10) are there capstone projects
no
(11) are there internships (maybe along with capstone projects)
no
(12) how applied? how theoretical?
thoughtful

University of Chicago

[Chicago] Masters of Science in Analytics
https://grahamschool.uchicago.edu/academic-programs/masters-degrees/analytics
This centralized, standalone program has a specific business focus and a goal of placing graduates into industry positions. Compared to other programs, there is a moderate amount of programming and math involved. All of the courses are created specifically for this program. Most students have at least three years of work experience before admission. A three-quarter-long capstone project is involved which often involves a problem taken from industry.
(1) Is it a standalone program or a track within another department/program or is it intended to
be connected with another major/program?

   It is a standalone program that usually takes 1.5 years to complete.
(2) Does the program have a specific focus (e.g., business analytics, bio data) or more general?

   MScA is for business professionals who aim to become analytics professionals with
advanced data analysis and modeling skills. Graduates of this program are employed for such
roles as data scientist, data analyst, business analytics professional, or manager of analytics
teams.
(3) Does the program have multiple tracks, emphasis areas, and if so what - are there different
threads?

   No, the program does not have tracks.
(4) Do the programs assume a secondary major / training in another discipline, how much is
disciplinary knowledge an important part of the training (vs. a more general approach)?

   No secondary major or training in another discipline is required.
(5) Does the program have more of a computational, mathematical, statistical emphasis?

   This program appears to have an even emphasis on all three.
(6) What kind of math and programming does it require (both as part of the program and in
terms of admission/entry to the program?)

   Fourteen courses are required (part or full time) plus a final capstone. This includes: one
Statistics Bootcamp, two Foundational Courses, seven Core Courses, one Business Strategy /
Project Management Course, three Electives, and a Capstone project. As part of the program:
machine learning and predictive analysis, statistical analysis, linear algebra & matrix
analysis, data mining principles, programming for analytics, database design &
implementation, marketing analytics, time series analysis & forecasting, linear & nonlinear
models, advanced python for streaming analytics, real time analytics, leadership skills, big
data and text analytics, research design for business applications, advanced machine learning,
programming for analytics, digital marketing analysis, data visualization, statistical analysis
review, optimization & simulation methods, research design for business applications. Some
of these classes are foundational, core, and elective.
(7) For grad programs, what kinds of students does it target for recruitment (undergrad majors,
work experience), and how much undergrad training in math, statistics, programming does it
expect?

   The average age is 32; 12% have 0-2 years work experience, 32% have 3-5 years, 26% have
6-10 years, 26% have 10+ years work experience, and 66% are male.
(8) What special courses were created for the major/minor/certificate - are there special math,
stats, and programming classes, what are the special data science courses (vs. repackaging
existing courses, or using cross-listing, to create a new interdisciplinary program?)

   Although some courses look to have conventional titles (e.g. linear & nonlinear models), they
are not as one might have expected taught by tenure track faculty in a conventionally-
associated department at U. of Chicago (e.g. statistics). Rather, all of the courses in this
program appear to be taught by data science professionals who work in industry but also do
teaching on the side (see https://grahamschool.uchicago.edu/academic-programs/masters-
degrees/analytics/instructors). All of the courses appear to be created specifically for this
program.
(9) How centralized is the center/program in which the program resides?
there is a centralized cohort approach

(10) are there capstone projects?
Yes, capstone projects take 3 quarters to complete and students can begin work on them after they have completed 5 courses. Students work in teams of 2-3 on capstone projects with a faculty advisor. Teams may identify a company with both a business problem and the resources needed to carry out the research process or work with analytics industry research partners. Results are presented in oral and written form and oral presentations are judged by a panel and a “best in showcase” is selected.

(11) are there internships (maybe along with capstone projects)?
No, there are not internships.

(12) how applied? how theoretical?
The orientation is more towards theoretical.

(13) other?
A unique feature is a leadership skills class (required) that teaches students how to make ties between data analysis and business objectives.

**[Chicago] Online Business Analytics Certificate**
https://grahamschool.uchicago.edu/academic-programs/professional-development/business-analytics
This centralized, four-course non-degree online certificate program exclusively focuses on business analytics and business intelligence. Compared to other programs, it is more applied, with a computational emphasis. Students interact online synchronously once per week.

(1) is it a standalone program or a track within another department/program or is it intended to be connected with another major/program?
It is a non-degree online certificate program; students do not earn credit toward a graduate degree.

(2) does the program have a specific focus (e.g., business analytics, bio data) or more general?
It has a specific focus. The Business Analytics certificate consists of two components: business intelligence and business analytics. The business intelligence courses develop students’ ability to work with data to evaluate and understand what has and is happening within the business. The business analytics courses develop students’ analytical skills to drive impactful insights from data by using standard statistical software tools.

(3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads?
There are not multiple tracks. There are only 4 courses required in total.

(4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)?
The assumption is that students have background in business.

(5) does the program have more of a computational, mathematical, statistical emphasis?
The emphasis appears to be more computational.

(6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)?
No specific math and programming background is mentioned.
(7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect?

No necessary prior training was mentioned.

(8) what special courses were created for the major/minor/certificate? Are there special math, stats, and programming classes? What are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)?

The Business Analytics certificate consists of four required online courses, all of which were developed specially for this program. They must be completed in a particular order. The first two are Business Intelligence Courses and the second two are Business Analytics Courses: (1) Exploring Data to Evaluate Business Practices: Databases and Reporting; (2) Exploring Data to Evaluate Business Practices: Exploratory Data Analysis and Visualization; (3) Data Analysis for Evidence Based Decision Making; (4) Data Mining for Evidence Based Decision Making. The courses include online course content, activities, and assessments. In addition, classes will meet virtually in a weekly one-hour synchronous session where students interact with the course instructor and each other. This combination of content accessible anytime combined with weekly synchronous sessions provides flexibility, while ensuring that students enjoy the benefits of interactive learning experience.

(9) how centralized is the center/program in which the program resides?

It is centralized; all courses and career counseling are all provided within the program, not spread across departments

(10) are there capstone projects?

There are not capstone projects.

(11) are there internships (maybe along with capstone projects)?

There are not internships.

(12) how applied? How theoretical?

This program is more applied.

[Chicago] Executive Program in Applied Data Analytics
https://coleridgeinitiative.org/training

This 12-day program is called the “Coleridge Initiative” and it is decentralized – jointly administered by the University of Chicago, NYU, and the University of Maryland. This program requires a quantitative masters degree or two years of work experience. All courses were specially created for the program.

(1) is it a standalone program or a track within another department/program or is it intended to be connected with another major/program?

It spans three universities but is not connected with a particular major or department.

(2) does the program have a specific focus (e.g., business analytics, bio data) or more general?

This program’s specific focus is “designed to address public sector challenges, but is applicable to a broad range of fields such as sociology, public health, computer science, survey statistics, and federal departments.”

(3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads?

This program does not have multiple tracks.
(4) do the programs assume a secondary major / training in another discipline? How much is disciplinary knowledge an important part of the training (vs. a more general approach)?

Yes, background is assumed: “this program is designed for professionals with a masters’ degree or above with some emphasis in a quantitative field (e.g. statistics, economics or computer science) or at least two years working in a hands-on, data-oriented field.”

(5) does the program have more of a computational, mathematical, statistical emphasis?

This program does not seem to emphasize one over the others.

(6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)?

They do not give specific requirements in terms of prior courses but they generally want either a previous masters degree in statistics, economics, or computer science, or they want 2 years work experience in one of these fields.

(7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect?

See (4) above.

(8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)?

All courses appear to be specially created for the program. The general format is 12 training days grouped into 4 modules of 3 days each. There is a primary focus on a different topic each day, as outlined below. In Module 1 (three days) there is: Intro and project scoping, Databases, SQL, and data exploration, and Intro to python for data analysis. In Module 2 (three days) there is: Python and APIs, Record Linkage, Network Analysis. In Module 3 there is: Intro to Machine Learning, Machine Learning, and Text Analytics. In Module 4 there is Web-scraping & Big Data lecture, Inference, Privacy and Confidentiality. Slack is one method of communication for the Applied Data Analytics program.

(9) how centralized is the center/program in which the program resides?

It is uncentralized; it spans three universities. See (1) above.

(10) are there capstone projects?

This is a very short program (12 training days total) but nonetheless they have an abbreviated version of a capstone experience, described as: “participants are grouped in teams of up to five members to work through an analytic project that answers a specific research question formulated by the team.”

(11) are there internships (maybe along with capstone projects)?

There are no internships.

(12) how applied? how theoretical?

This program is in the middle.

University of Virginia

[Virginia] Master’s in Data Science
https://dsi.virginia.edu/degrees/msds-degree

This general 11-month masters program has a degree granted by an institute, rather than a department. Most students come straight out of college, and most are from UVA. A moderate
amount of programming and math are required as background. Students work in teams throughout the duration of the program. For the capstone project each team is paired with (and funded by) an industry professional. Several large, complicated data sets are woven across courses to increase the program’s cohesion. There are also dual-degree and executive options. Prospective “sponsor” companies can contact the program via an online link to request being a sponsor. This program has a dedicated career counselor and a dedicated stipend to cover professional expenses related to job interviewing; starting salary for graduates averages 90K.

1) is it a standalone program or a track within another department/program or is it intended to be connected with another major/program?
   The masters in data science is a standalone 11-month program, and the degree is granted by an institute (Data Science Institute), not a department.

2) does the program have a specific focus (e.g., business analytics, bio data) or more general?
   This program is general.

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads?
   No multiple tracks were listed. Student only have approximately 2 electives total.

4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)?
   The regular Masters in Data Science uses a general approach: any 3-year or 4-year undergraduate degree is acceptable. There is also a Dual Degree option: Masters in Data Science + MBA during 2 years or Masters in Data Science + MD during 2 years or Masters in Data Science + PhD in Nursing. There is also an Executive option (credit or noncredit, customizable content, flexible formats, varied learning environments)

5) does the program have more of a computational, mathematical, statistical emphasis?
   Not mentioned and the program did not publish their curriculum online. Alumni reviewers mentioned it was a “nice balance of stats and math.”

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)?
   Pre-requisites prior to entry into the program are: Single variable calculus (similar to UVA MATH 1210-1220, MATH 1310-1320 or APMA 1090-1110), Linear algebra or matrix algebra (similar to UVA MATH 3350, MATH 3351 or APMA 3080), Introductory statistics (similar to UVA STAT 2020 or STAT 2120), and Introductory programming (similar to UVA CS 1110-1112).
   As part of the program, although the curriculum was not posted online, alumni reviewers stated that: The course started with a summer semester (6 weeks) in which built programming skills in R, python and SAS. And then the fall semester (6 weeks) involved: Linear Modelling and Regression in Stats and Data Mining where they were introduced to various algorithms that will be useful in this field. Then in the January term (very short : 8 days) they covered ethical and privacy issues. In the spring semester (6 weeks) they have courses built on the skills that they developed thus far: Data Mining, Machine learning and a paid capstone project. Electives: Reinforcement Learning, Computer Vision and Language, Data Visualisation, Cloud Computing, Databases, Information Retrieval.
   Alumni said the programming classes were review for CS majors and the Linear Modeling class was review for the stats majors but that everyone learned from the data mining and machine learning courses.
7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect?

No work experience is necessary. A 3- or 4-year undergraduate degree plus the 4 pre-requisite courses listed in 6) above are required. According to alumni who posted reviews of the program online: cohorts are around 47 students from a pool of 400 applicants, most fresh from undergrad, and most UVA grads. Mostly CS, Statistics, or Economics majors but include English, Russian, and Physics majors.

8) what special courses were created for the major/minor/certificate? Are there special math, stats, and programming classes? What are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)?

The curriculum was not published online, but alumni postings indicated that at least the data mining and machine learning and capstone courses were specially created for the masters program. See 6) regarding courses.

9) how centralized is the center/program in which the program resides?

This program is very centralized; 80% of courses are shared among the cohort.

Several large, complicated data sets are woven across courses to increase the program’s cohesion.

Students consistently work together in teams (even for the capstone project), building strong relationships and practical skills in the process.

10) are there capstone projects?

Yes; project teams consist of two to four MSDS students and one or more faculty advisors. These teams work with sponsoring organizations to provide valuable recommendations to address strategic and operational issues. Depending on the needs of the sponsor, teams may develop web-based applications that can support ongoing decision making. Note: Industry “sponsors” provide financial backing (e.g., funding for commercial computing needs), data, and research questions, and they meet regularly with students and faculty advisor. Students appear to be paid for their capstone work by the sponsor. In return, the products of the capstone need to include a conference presentation and publishable paper. Prospective “sponsor” companies can contact the program via an online link to request being a sponsor. Students rank their interests in various sponsor projects but once assigned cannot switch.

11) are there internships (maybe along with capstone projects)?

There are no internships, just the capstone.

12) how applied? how theoretical?

This program is in the middle.

13) other?

They have a dedicated career counselor and a dedicated stipend to cover professional expenses related to job interviewing. In last year’s cohort all but two ended had jobs immediately after graduation and those two had reasons for not getting jobs right away.

Average salary was $89,000.

[Virginia] Business Analytics Track within M.S. in Commerce
https://www.commerce.virginia.edu/ms-commerce/business-analytics

This is not a standalone program but rather a track focusing on business analytics within the M.S. in Commerce; students are recruited from the masters program in Commerce. The focus of the program is more statistical than mathematical or computational.
(1) is it a standalone program or a track within another department/program or is it intended to be connected with another major/program?
   This is a track – specifically, a 15 Credit hour track on “business analytics” within the M.S. in Commerce in the School of Commerce
(2) does the program have a specific focus (e.g., business analytics, bio data) or more general?
   The specific focus of this program is on business analytics.
(3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads?
   This program does not have multiple tracks within the Business analytics track.
(4) do the programs assume a secondary major / training in another discipline? How much is disciplinary knowledge an important part of the training (vs. a more general approach)?
   Training in another discipline is assumed; this program assumes you are getting the MS in Commerce
(5) does the program have more of a computational, mathematical, statistical emphasis?
   The focus of the program is more statistical than mathematical or computational.
(6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)?
   Requires graduating within 18 months of matriculation with a strong liberal arts, science or engineering undergraduate degree (not a business major). Required pre-requisites are Introduction to Financial Accounting, Principles of Microeconomics and Introduction to Statistics.
(7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect?
   See (6)
(8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)?
   Courses specially created for the program are: Intro to Business Analytics, Customer Analytics, and Web Analytics for e-commerce. Courses or topics repackaged and repurposed for this program are: advanced quantitative analysis, logistic regression, discriminant and cluster analysis, factor analysis, regression models, structural equation models, ANOVA. (seems to be a repackaged course)
(9) how centralized is the center/program in which the program resides?
   This program is centralized within the School of Commerce.
(10) are there capstone projects?
    There are no capstone projects.
(11) are there internships (maybe along with capstone projects)?
    There are no internships.
(12) how applied? how theoretical?
    This program is more applied.
Brown University

**[Brown] Pre-College**
Summer@Brown offers courses that are designed for students looking to experience college-level academics on an Ivy League campus. Reflecting Brown’s broad liberal arts curriculum, Summer@Brown features nearly 200 courses, designed and taught at the level of first-year college courses, and ranging in length from one to four weeks. Programs include: Biomedical Informatics and Data Science for Biomedicine and Health Care, Visualizing Your Data: Graphical Programming in R, A Data-Centric Introduction to Programming

**[Brown] Undergraduate in Applied Mathematics**
1) *is it a standalone program or a track within another department/program - is it intended to be connected with another major/program.*
   A standalone major (concentration). The standard Applied Mathematics concentrations lead to either the A.B. or Sc.B. degrees. The program is very flexible. Numerous joint programs with other departments are described below, and individual concentrations suited to particular needs can be arranged. The range of offerings, either within Applied Mathematics alone or in combination with offerings of other departments, provides almost endless opportunities. There are three official joint honors programs with other departments (economics, biology, computer science).
2) *does the program have a specific focus (e.g., business analytics, bio data) or more general*
   General focus, although joint concentrations have specific focus
3) *does the program have multiple tracks, emphasis areas, and if so what - are there different threads*
   The major is very flexible, with guidance on how to develop a course of study if you want to go to graduate school, if you want to go into accounting or insurance, if you want to focus on computation, etc. Within this flexibility, there are planned joint concentrations in economics, biology, and computer science.
4) *do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)*
   Depends on specific track
5) *does the program have more of a computational, mathematical, statistical emphasis*
   Highly varied, depends on the specific course sequence
6) *what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)*
   Core program includes: Calculus sequence: Math 0090, 0100, 0180 (or 0350); Linear Algebra: MATH 0520 or preferably 0540; Differential Equations: APMA 0350 and 0360; Computing Course: APMA 0090 or 0160 or CSCI 0040, 0150, or 0170
7) *for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect*
   N/A
8) *what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)*

74
All courses are housed within the applied mathematics department, although the program is flexible. Further, the joint concentrations have required courses in the applied science department and their collaborating departments (i.e., biology, computer science). APMA 0110. What’s the Big Deal with Data Science? : This seminar serves as a practical introduction to the interdisciplinary field of data science. Over the course of the semester, students will be exposed to the diversity of questions that data science can address by reading current scholarly works from leading researchers. Through hands-on labs and experiences, students will gain facility with computational and visualization techniques for uncovering meaning from large numerical and text-based data sets. Ultimately, students will gain fluency with data science vocabulary and ideas. There are no prerequisites for this course.

9) how centralized is the center/program in which the program resides
   Applied mathematics is a separate department

10) are there capstone projects
   The biology and computer science joint concentrations have senior capstone requirements where students work with a faculty member.

11) are there internships (maybe along with capstone projects)
   Opportunities listed, but not required. Website indicates many faculty like working with undergrad AM students during the summer

12) how applied? how theoretical?
   Highly varied

[Brown] Masters of Science in Biostats-Health Data Science track
https://www.brown.edu/academics/public-health/biostatistics/health-data-science-track

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program.
   This is a track within another program (biostats masters)

2) does the program have a specific focus (e.g., business analytics, bio data) or more general health science

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
   This is a track within a larger program

4) does the program assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
   Program is interdisciplinary of statistics, mathematics, computer science, and the application field

5) does the program have more of a computational, mathematical, statistical emphasis mathematical

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program
   Students should have good mathematical preparation, including multivariable calculus, linear algebra, and mathematical probability using calculus. You may apply before completing these courses, however they need to have been completed prior to starting the program.

7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
Most courses use R for programming. We do not require a specific type of undergraduate degree. Applications from students in applied fields such as biology, biochemistry, economics and computer science are strongly encouraged, with the understanding that necessary mathematical coursework needs to be completed before enrollment in the program.

8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)

Nearly all classes in biostats, although this track seems to have only one specifically created course in health data science

9) how centralized is the center/program in which the program resides

Resides in the biostats department, which is strongly affiliated with Center for Statistical Sciences

10) are there capstone projects

There are options for AM or ScM. The ScM requires a research project.

11) are there internships (maybe along with capstone projects)

No

12) how applied? how theoretical?

Seems more theoretical, with some application

13) other

Began in 2016; also available as a 5th year program

[Brown] Masters in Data Science
https://www.brown.edu/initiatives/data-science/academic-programs/masters-data-science

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program.

standalone

2) does the program have a specific focus (e.g., business analytics, bio data) or more general

general

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads

No tracks

4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)

Rooted in a research collaboration among four very strong academic departments (Applied Mathematics, Biostatistics, Computer Science, and Mathematics), the master's program will offer a rigorous, distinctive, and attractive education for people building careers in Data Science and/or in Big Data Management.

5) does the program have more of a computational, mathematical, statistical emphasis

computational

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program

The program will be conducted over one academic year plus one summer, with the option for an additional pre-program summer for students who lack one or more of the basic prerequisites.
3 credits in mathematical and statistical foundations, 3 credits in data and computational science, 1 credit in societal implications and opportunities, 1 elective credit to be drawn from a wide range of focused applications or deeper theoretical exploration, and 1 credit capstone experience.

7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect

Students entering the program will be required to have completed at least a year of calculus, a semester of linear algebra, a semester of calculus-based probability and statistics, and an introduction to programming. We also admit exceptional students who lack one or more of the minimum requirements in linear algebra, probability and statistics, and computer science. The four departments (Math, Applied Math, Computer Science, and Biostatistics) will offer a suitable course in each of these three topics during the Brown summer session before the first semester.

8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)

Newly created courses

9) how centralized is the center/program in which the program resides

Affiliated with new data science initiative https://www.brown.edu/initiatives/data-science/, collaboration among 4 departments

10) are there capstone projects

Yes. The regular program includes two semesters of coursework and a one-summer (5-10 week) capstone project focused on data analysis in a particular application area.

11) are there internships (maybe along with capstone projects)

No, although describes strong industry relationships of faculty

12) how applied? how theoretical?

Very applied

13) other

Also available as a 5th year program

New York University

[NYU] MS in Business Analytics (NYU Stern)

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program

MSBA is a one year long, part-time program divided into five on-site class sessions that take place at NYU Stern and in two rotating global locations.

2) does the program have a specific focus (e.g., business analytics, bio data) or more general Business Analytics

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads

No
4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)

“The MS in Business Analytics Program is designed for experienced managers with oversight for and/or increasing responsibilities in areas that help drive greater efficiencies, manage costs, and identify growth opportunities. Participants come from a broad range of sectors: consumer products, energy, financial, healthcare, pharmaceutical, industrials, technologies, NGOs, consulting, etc”

5) does the program have more of a computational, mathematical, statistical emphasis

Yes: Statistics Using R

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)

Primarily use of R

7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect

“The MS in Business Analytics Program is designed for experienced managers with oversight for and/or increasing responsibilities in areas that help drive greater efficiencies, manage costs, and identify growth opportunities. Participants come from a broad range of sectors: consumer products, energy, financial, healthcare, pharmaceutical, industrials, technologies, NGOs, consulting, etc”

8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)

Entire curriculum appears to be specialized to this degree (part-time MSBA), has very clear sequencing to the degree curriculum: Foundations of Statistics Using R, Digital Marketing Analytics, Data Science for Business Analytics, Dealing with Data, Data Mining in R, Data Driven Decision Making, Operations Analytics, Data Visualization, Network Analytics, Decision Under Risk, Revenue Management and Pricing, Data Privacy and Ethics, Strategy, Change and Analytics, Strategic Capstone

9) how centralized is the center/program in which the program resides

highly

10) are there capstone projects

yes

11) are there internships (maybe along with capstone projects)

no

12) how applied? how theoretical?

Entirely applied

[NYU] MBA with a Specialization in Business Analytics (NYU Stern)

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program

Inside regular MBA

2) does the program have a specific focus (e.g., business analytics, bio data) or more general Business Analytics

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
courses from ECON, INFO, MKTG, OPMG, STAT

4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
   Unclear

5) does the program have more of a computational, mathematical, statistical emphasis
   Depends see above

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
   Depends

7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
   MBA students

8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
   unclear

9) how centralized is the center/program in which the program resides
   Diffuse

10) are there capstone projects
    no

11) are there internships (maybe along with capstone projects)
    unknown

12) how applied? how theoretical?
    both applied and theoretical

[NYU] MA Applied Quantitative Research (NYU Arts & Science: Sociology)

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
   Sociology Department

2) does the program have a specific focus (e.g., business analytics, bio data) or more general Quantitative Social Research

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads

4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
   Unclear

5) does the program have more of a computational, mathematical, statistical emphasis
   Statistical

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
   Unavailable online
7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
   Unclear
8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
   Design of Social Research, Applied Quantitative Analysis I & II, Data Analysis, AQR Workshops
9) how centralized is the center/program in which the program resides
   Modestly centralized
10) are there capstone projects
    yes, MA Thesis
11) are there internships (maybe along with capstone projects)
    Unclear
12) how applied? how theoretical?
    Very applied

[NYU] Master’s in Data Science
https://cds.nyu.edu/academics/ms-in-data-science/
MS in Data Science (NYU Center for Data Science) (claims to be the first in the country)

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
   Standalone program in Center for Data Science. 36 credits, 2 years. http://gsas.nyu.edu/bulletin/data-science/ms-data-science.html
2) does the program have a specific focus (e.g., business analytics, bio data) or more general
   General approach to Data Science as its own discipline: “The Master of Science in Data Science is a highly-selective program for students with a strong background in mathematics, computer science, and applied statistics. The degree focuses on the development of new methods for data science…. A new discipline has emerged to address the need for professionals and researchers to deal with the “data tidal wave.” Its object is to provide the underlying theory and methods of the data revolution. This emergent discipline is known by several names. We call it “data science,” and we have created the world’s first MS degree program devoted to it.”
3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
   Yes: Data Science Track, Data Science Big Data Track, Data Science Mathematics and Data Track, Data Science Natural Language Processing Track, Data Science Physics Track, Data Science Biology Track (This track will be available for incoming Fall 2018 students.)
4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
   Unclear: “The MS in Business Analytics Program is designed for experienced managers with oversight for and/or increasing responsibilities in areas that help drive greater efficiencies, manage costs, and identify growth opportunities. Participants come from a broad range of
sectors: consumer products, energy, financial, healthcare, pharmaceutical, industrials, technologies, NGOs, consulting, etc”

5) does the program have more of a computational, mathematical, statistical emphasis

Computational and Statistical emphasis

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)

“Programming languages are decided by the professor of each course. However, over the past few years, Python has been widely used.”

7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect

Successful applicants to the MSDS come from many different undergraduate backgrounds, including degrees in Statistics, Computer Science, Mathematics, Engineering, Economics, Business, Biology, Physics and Psychology. Many of our students join us directly from undergraduate, but we also very much welcome evidence of relevant work experience – and clear employment goals once the MSDS is completed – in data science. Past experience and career aspiration goals can be related to commercial industry, government, academia or some other sector. Most of the required courses are offered in the evenings. Elective courses are offered both during the day and in the evenings. 5-year window to complete 2 years of full time course work.

Prerequisites: https://cds.nyu.edu/admissions/ms-requirements/ To be considered for the program, you will be required to have completed the following (or equivalents): Calculus I: limits, derivatives, series, integrals, etc.; Linear Algebra; Intro to Computer Science (or an equivalent “CS-101” programming course): We have no set requirements as regards specific languages, but we generally expect serious academic and/or professional experience with Python and R at a minimum; One of Calculus II, Probability, Statistics, or an advanced physics, engineering, or econometrics course with heavy mathematical content; Preference is given to applicants with prior exposure to machine learning, computational statistics, data mining, large-scale scientific computing, operations research (either in an academic or professional context), as well as to applicants with significantly more mathematical and/or computer science training than the minimum requirements listed above.

8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)

Introduction to Data Science; Probability and Statistics for Data Science; Machine Learning and Computational Statistics; Big Data; Inference and Representation; Deep Learning; Natural Language Processing with Representation Learning; Natural Language Understanding and Computational Semantics; Optimization-based Data Analysis

9) how centralized is the center/program in which the program resides

Centralized but also draws on numerous electives outside the center

10) are there capstone projects

Yes

11) are there internships (maybe along with capstone projects)

Yes, “CPT and OPT are options for Data Science students. CPT is “curricular practical training,” a program which allows international students to obtain paid internships for credit
after the first year of the degree. OPT is “optional practical training,” a program which allows international students to obtain paid internships without credit either during or after the degree.”

12) how applied? how theoretical?
Primarily applied, but approaches DS as its own discipline

13) Other
Program is evolving rapidly, requirements have changed from 2017 to 2018 and website promises new electives for fall 2018.

[NYU] MS in Scientific Computing (Data Science Concentration) (NYU Courant Institute of Mathematical Sciences)

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
“The departments of mathematics and computer science at NYU’s Courant Institute of Mathematical Sciences offer a master's degree in scientific computing.”

2) does the program have a specific focus (e.g., business analytics, bio data) or more general
“The program provides broad yet rigorous training in areas of mathematics and computer science related to scientific computing. It aims to prepare people with the right talents and background for a technical career doing practical computing... The program accommodates both full-time and part-time students, with most courses meeting in the evening. The masters program focuses on computational science, which includes modeling and numerical simulation as used in engineering design, development, and optimization. While data science is an increasingly important aspect of computational science, this program is distinct and different from the recently-created Masters of Science in Data Science within the NYU Center for Data Science. Students specifically interested in data science are encouraged to apply to that program instead.”

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads

4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
Important see requirements below

5) does the program have more of a computational, mathematical, statistical emphasis computational, mathematical

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
The program requires at least three semesters of Calculus (including multivariate calculus), as well as linear algebra. Experience with programming in a high-level language (e.g., Java, C, C++, Fortran, Python) as well as data structures and algorithms, equivalent to a first-year sequence in computer science, is also required. It is highly desirable that applicants have undergraduate major or significant experience in mathematics, a quantitative science or engineering, or economics.

7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect

It aims to prepare people with the right talents and background for a technical career doing practical computing. The program accommodates both full-time and part-time students, with most courses meeting in the evening. The program is self-contained and terminal, providing a complete set of skills in a field where the need is greater than the supply. The program focuses on the mathematics and computer science related to advanced computer modeling and simulation, and is similar in structure to terminal master’s programs in engineering, combining classroom training with practical experience.

8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)

unclear

9) how centralized is the center/program in which the program resides

modestly centralized

10) are there capstone projects

yes, capstone

11) are there internships (maybe along with capstone projects)

no

12) how applied? how theoretical?

very applied

[NYU] PhD in Computer Science, Machine Learning/AI specialization (NYU Courant Institute of Mathematical Sciences)

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program

The department of computer science at NYU's Courant Institute of Mathematical Sciences

2) does the program have a specific focus (e.g., business analytics, bio data) or more general Machine Learning/AI specialization within CS PhD

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads

Machine Learning/AI specialization

4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)

Very important, breadth and depth prerequisites

5) does the program have more of a computational, mathematical, statistical emphasis computational

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
Algorithms, Systems, Applications, Major Exams
7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
   Significant previous work in CS
8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
   unclear
9) how centralized is the center/program in which the program resides
   very centralized
10) are there capstone projects
    yes, thesis
11) are there internships (maybe along with capstone projects)
    unclear
12) how applied? how theoretical?
    very theoretical

**MS in Applied Statistics for Social Science Research (NYU Steinhardt School of Culture, Education, and Human Development)**
1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
   The MS program is an initiative of the NYU Steinhardt Center for the Promotion of Research Involving Innovative Statistical Methodology (PRIISM) and is an integrated part of the larger university-wide initiative in Data Science and Statistics, which offers several new master’s degrees, including this MS program.
2) does the program have a specific focus (e.g., business analytics, bio data) or more general Applied emphasis on statistical study of social change as tracked by policy and advocacy organizations and in fields including public health, politics, psychology and psychometrics, demography, sociology, labor markets, and education.
3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
   no
4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
   No, but those with statistical background can complete degree in 34 credit instead of 42-47
5) does the program have more of a computational, mathematical, statistical emphasis
   statistical
6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
   unclear
7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
   none
8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)

range of courses

9) how centralized is the center/program in which the program resides

mixed core and electives (many outside the program)

10) are there capstone projects

yes

11) are there internships (maybe along with capstone projects)

yes

12) how applied? how theoretical?

very applied

[NYU] Ph.D. in Computer Science with focus on Big Data Management, Analysis, and Visualization (NYU Tandon School of Engineering)

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program

Tracks in Big Data Management, Analysis, and Visualization in the Department of Computer Science and Engineering

2) does the program have a specific focus (e.g., business analytics, bio data) or more general

Tracks in Big Data Management, Analysis, and Visualization

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads

Tracks in Big Data Management, Analysis, and Visualization

4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)

CS focused

5) does the program have more of a computational, mathematical, statistical emphasis

computational

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)

satisfy a breadth course requirement, intended to ensure broad knowledge of computer science,
satisfy a depth requirement, consisting of an oral qualifying exam presentation with a written report, to ensure the student's ability to do research

7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect

unclear

8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)

unclear

9) how centralized is the center/program in which the program resides

centralized

10) are there capstone projects
yes, PhD thesis

11) are there internships (maybe along with capstone projects)
   no, unclear

12) how applied? how theoretical?
   very theoretical

**University of Rochester**

[Rochester] BA, BS in Data Science (Goergen Institute for Data Science)

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
   Stand alone program in the Goergen Institute for Data Science

2) does the program have a specific focus (e.g., business analytics, bio data) or more general
   More general

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
   Yes: Biology, Brain and cognitive sciences, Computer science, statistics, and mathematic, Earth and environmental science, Physics, Economics and business, Political science

4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
   No

5) does the program have more of a computational, mathematical, statistical emphasis
   Unclear, probably computations and mathematical

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
   Plus one of the following sequences: MTH 161: Calculus I and MTH 162: Calculus II; MTH 141: Calculus I, MTH 142: Calculus II, and MTH 143: Calculus III; MTH 171: Calculus I (Honors) and MTH 172: Calculus II (Honors)
   Core: One of the following: MTH 165: Linear Algebra with Differential Equations; MTH 173: Calculus III (Honors); Plus one of the following: DSC/CSC 262: Computational Introduction to Statistics; STT 213: Elements of Probability and Mathematical Statistics; STT 212: Applied Statistics for the Biological and Physical Sciences I
   Plus one of the following: DSC/CSC 265: Intermediate Statistical and Computational Methods; Both STT 216: Applied Statistics II and STT 226W: Introduction to Linear Models
   Plus all of the following: CSC 240: Data Mining; CSC 242: Introduction to Artificial Intelligence; CSC 261: Database Systems; CSC 282: Design and Analysis of Efficient Algorithms; DSC 383W: Data Science Capstone (fall semester of senior year)

7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
   N/A
8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
   DSC 383W: Data Science Capstone (fall semester of senior year)
9) how centralized is the center/program in which the program resides
   not very
10) are there capstone projects
    yes, DSC 383W: Data Science Capstone (fall semester of senior year)
11) are there internships (maybe along with capstone projects)
    maybe
12) how applied? how theoretical?
    applied

[Rochester] MS in Data Science (Goergen Institute for Data Science)
1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
   stand-alone program in the Goergen Institute for Data Science
2) does the program have a specific focus (e.g., business analytics, bio data) or more general
   Computational and statistical methods, Health and biomedical sciences, Business and social science
3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
   Computational and statistical methods, Health and biomedical sciences, Business and social science
4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
   An optional summer bridging course for students who come without a strong computer science background.
5) does the program have more of a computational, mathematical, statistical emphasis
   all three
6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
   unclear
7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
   unclear
8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
   Core Courses: DSC 462: Computational Introduction to Statistics; DSC 465: Intermediate Statistical and Computational Methods; CSC 440: Data Mining (offered fall and spring)
   CSC 461: Database Systems (offered fall and spring)
9) how centralized is the center/program in which the program resides
   centralized
10) are there capstone projects
Yes, a required 4 credit practicum in which the student works in a team to implement a significant system or analysis with a final oral presentation provided by each student. A committee of two faculty members from within the institute will evaluate the final oral presentation in order for it to serve as the master’s degree exit exam.

11) Are there internships (maybe along with capstone projects)?
unclear

12) How applied? How theoretical?
applied

**University of Washington**

**[Washington] Undergraduate: Data Science Option**
http://escience.washington.edu/education/undergraduate/

1) Is it a standalone program or a track within another department/program - is it intended to be connected with another major/program.
   Connected. This is an “add on” to existing undergraduate programs

2) Does the program have a specific focus (e.g., business analytics, bio data) or more general
   general

3) Does the program have multiple tracks, emphasis areas, and if so what - are there different threads?
   No, single core curriculum, although students have a primary undergraduate program

4) Do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
   Yes, all are in a primary undergraduate program. This is an “add on”. This option exists in applied computational and mathematical sciences, computer science and engineering, and information school

5) Does the program have more of a computational, mathematical, statistical emphasis?
   computational

6) What kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program)
   unclear

7) For grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
   based on requirements of each affiliated program

8) What special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
   At the undergraduate level, a student who will earn a data science specialization must complete courses in the following areas: Required areas with recommended courses: Programming: e.g. CSE163 or CSE143; Machine learning: e.g. CSE416/STAT416, STAT435, INFO 371; Societal implications of data science: e.g. SOC 201: Special Topics (Topic: Data and Society);
Required to cover at least two areas: Data management: e.g. CSE414 or INFO 445; Data visualization and communication: e.g. CSE412, INFO474, or HCDE411; Advanced statistics and probability: Department-specific course choices
Optional: Introduction to data science: e.g. STAT180/CSE180/INFO180

9) how centralized is the center/program in which the program resides
decentralized, affiliated with eScience institute

10) are there capstone projects
   no

11) are there internships (maybe along with capstone projects)
   no

12) how applied? how theoretical?
   applied

[Washington] Master of science in Data Science
https://www.datasciencemasters.uw.edu/

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program.
   standalone

2) does the program have a specific focus (e.g., business analytics, bio data) or more general
   general

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
   no tracks

4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
   Our interdisciplinary curriculum was developed by leading faculty from six top-ranked departments and schools at the UW (applied math, biostats, school of computer science and engineering, school of human centered design and engineering, information school, department of statistics)

5) does the program have more of a computational, mathematical, statistical emphasis
   Computational. Many of the courses will emphasize team-based data analysis and engineering work and will involve working in small groups to complete one or more guided practicum projects per quarter.

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program
   unknown

7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
   Depending on the course, students can expect an emphasis on Python and R programming, and some assignment work with Java; Admission to the UW Master of Science in Data Science program is competitive. To be successful in the program, students need a solid background in mathematics (three-class calculus sequence and a linear algebra course), computer programming (two-course introductory sequence in programming) and communication
8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)

Mostly newly created courses, although one is cross-listed


9) how centralized is the center/program in which the program resides

Organized by the eScience Institute, program is collaboration of 6 departments. Program is run by an executive committee. Executive Committee members also serve on the Interdisciplinary Data Science Group (IDSG), a cross-departmental group that provides strategic direction for the master’s program. http://escience.washington.edu/

10) are there capstone projects

Yes. 2-quarter capstone experience. The final course is a capstone project where students get to solve a real-world data analysis challenge facing a local organization.

11) are there internships (maybe along with capstone projects)

Not as part of the program, although website has big focus on industry connections

12) how applied? how theoretical?

Very applied. Emphasizes industry connections and real-world data analysis

[Washington] Doctoral


1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program.

Connected, each associated department has an IGERT track. The IGERT big data PhD program is affiliated with 7 departments. Every student is associated with a primary department and complete doctoral coursework associated with that program. But, they also complete coursework associated with IGERT (which may overlap with their primary department in different ways). Each of the associated departments also has an Advanced Data Science Option, which overlaps with IGERT, but is not exactly the same. The department’s Advanced Data Science Option may have requirements in addition to IGERT. The core program of ADSO and IGERT are the same, although ADSO students don’t have a secondary advisor, interdisciplinary research project, or internship requirement.

2) does the program have a specific focus (e.g., business analytics, bio data) or more general

Varies by primary department

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads

Based on the associated departments

4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
Associated departments: astronomy, biology, chemical engineering, computer science and engineering, genome sciences, oceanography, and statistics

5) does the program have more of a computational, mathematical, statistical emphasis

   Mathematical

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program

   Substantial list of pre-reqs, including in programming, calculus and linear algebra, probability and statistics

7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect

   The Big Data IGERT is a highly selective 2-year Fellowship for PhD students interested in research in Big Data. Students have already been selected into a PhD program in one of the associated departments

8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program

   Core program: machine learning, data management, data visualization, statistics, data science seminar

   Courses are mostly cross-listed from the various departments. Students have a secondary advisor outside of their primary department. The role of this secondary advisor is to provide guidance on data science methods

9) how centralized is the center/program in which the program resides

   Decentralized. Must be affiliated with a primary department, and beyond core curriculum, how IGERT interacts with primary department varies. IGERT is organized by eScience initiative

10) are there capstone projects

   All IGERT students are expected to pursue a thesis in the area of Big Data / Data Science. In this context, the program expects all students to either contribute to the algorithmic, statistical or systems infrastructure for Big Data and release this work in the form of open-source software, or engage and utilize state-of-the-art Big Data tools to address a core scientific or engineering task.

   Interdisciplinary Project: Students are recommended to do a two-quarter long research project with a secondary advisor with complementary skills, e.g., a computer science student may do the project in an astronomy lab. This project will culminate in a presentation and a written report of a quality that the secondary advisor would consider publishable in a top venue in the complementary field. The program strongly encourages all students to present this work in the form of a paper or at least a poster at a conference or workshop.

11) are there internships (maybe along with capstone projects)

   Every student is recommended to do at least one internship in industry or National Labs during their PhD. These internships must cover practical work around Big Data.

12) how applied? how theoretical?

   more theoretical, although plenty of application

13) other

   It appears that IGERT is only accepting one more year of cohorts.
[Washington] Data Science Graduate Option
http://escience.washington.edu/education/phd/data-science-graduate-option/

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program.
   Connected. This is an “add on” to existing graduate (masters/doctoral) programs
2) does the program have a specific focus (e.g., business analytics, bio data) or more general
   general
3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
   No, single core curriculum, although students have a primary graduate program
4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
   Yes, all are in a primary graduate program. This is an “add on”. This option exists in the astronomy PhD program and the masters and PhD programs in chemical engineering
5) does the program have more of a computational, mathematical, statistical emphasis
   computational
6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program
   unclear
7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect
   based on requirements of each affiliated graduate program
8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)
   To ensure that the meaning of the Data Science Options remains consistent across departments and schools, all units in this program agree to shape their options by following the same overall framework. In this framework, to complete the “Data Science” option, students take at least three courses out of the following four areas, each course from a different area: software development for data science, statistics and machine learning, data management and visualization, department-specific data science courses, eScience community seminar; Uses existing courses from various departments
9) how centralized is the center/program in which the program resides
   decentralized, affiliated with eScience institute
10) are there capstone projects
    no
11) are there internships (maybe along with capstone projects)
    no
12) how applied? how theoretical?
    applied
[Washington] Professional Certification
https://www.pce.uw.edu/certificates/data-science

1) is it a standalone program or a track within another department/program - is it intended to be connected with another major/program
standalone

2) does the program have a specific focus (e.g., business analytics, bio data) or more general
general

3) does the program have multiple tracks, emphasis areas, and if so what - are there different threads
    no, single core curriculum

4) do the programs assume a secondary major / training in another discipline, how much is disciplinary knowledge an important part of the training (vs. a more general approach)
    no

5) does the program have more of a computational, mathematical, statistical emphasis
    computational

6) what kind of math and programming does it require (both as part of the program and in terms of admission/entry to the program
    Python

7) for grad programs, what kinds of students does it target for recruitment (undergrad majors, work experience), and how much undergrad training in math, statistics, programming does it expect

   Technically oriented professionals who work with data and have some background in statistics, SQL and programming. All applicants must take the data science qualifications assessment quiz. The assessment includes 30 multiple-choice questions divided into three sections: statistics and linear algebra, programming, and databases and SQL. Applicants must earn a total score of at least 18 out of 30, with a minimum score of 6 out of 10 in each section.
   Once you start the assessment, you must complete it within 90 minutes. After you submit your quiz, you’ll be able to see your total score (section scores are not available). Please use the same email address on both your assessment and your application.
   Applicants should also have some background in statistics, programming, and SQL including:
   At least one year of recent experience using a programming language, preferably Python, Java, R or C#; Fundamental understanding of databases and SQL; Basic conceptual knowledge and familiarity with statistics and linear algebra concepts such as vectors, matrices, eigenvectors and eigenvalues, simple matrix calculus (derivatives), distributions, quintiles, average and standard deviation as measures of distribution, and the concept of uncertain estimates of quantities as embodied by confidence limits or error bars. Applicants who need help meeting this requirement may consider enrolling in the course Foundations of Statistics.

8) what special courses were created for the major/minor/certificate - are there special math, stats, and programming classes, what are the special data science courses (vs. repackaging existing courses, or using cross-listing, to create a new interdisciplinary program)

   3 course sequence: Process and tools, methods for data analysis, machine learning techniques

9) how centralized is the center/program in which the program resides
Affiliated with eScience institute, but run by office of continuing education

10) are there capstone projects
   no

11) are there internships (maybe along with capstone projects)
   no

12) how applied? how theoretical?
   applied
APPENDIX G: Example Faculty Hires

This appendix lists descriptions for a number of example hires as a way to illustrate the range of possible areas and disciplines that could be part of the VDSI.

- An astronomer doing cutting-edge research on analyzing and visualizing the next generation of petabyte-scale astronomical surveys, like LSST.

- A visionary mathematician doing research in and developing algorithms for machine learning, sparse data, compressed sensing, random networks, multi-scale methods, applied harmonic analysis, geometric and topological methods in data analysis, and other foundational areas of data science.

- A legal or policy scholar focusing on designing public policy and legal institutional capacity for managing the impacts of AI in society (e.g., impacts on criminal justice, employment bias, financial system stability, etc.)

- An engineer with expertise in networked data, cyber-physical systems, and data analytics for autonomous vehicles integration.

- A scholar of higher education who harnesses data science techniques to analyze patterns and content of student and instructor interactions in online courses to better understand student engagement and connections to student learning.

- A finance scholar focusing on the use of data sciences to forge new understandings of complex financial systems (e.g., defining the drivers of systemic risk and indicia of imminent cascade failures).

- A legal or policy scholar focusing on the use of data sciences to forge new understandings of legal and policy institutions (e.g., using natural language processing to probe large text corpuses such as case law or voting records; developing predictive analytics for regulatory enforcement and compliance)

- A satellite remote sensing scholar focusing on mass satellite imagery analysis to model economic development decision-making at regional- to continental-scales.

- A computer scientist doing cutting-edge research on managing large-scale data stores, machine learning, or data mining, with applications that are synergistic with focus areas at Vanderbilt, such as medical imaging or clinical phenotype prediction.

- A scholar who bridges neuroscience/psychology and computer science, by developing neurally/psychologically plausible deep learning models of vision and/or cognition.

- A sociologist who applies data science methods to investigate large-scale social networks and their impact on emergency response situations.
• A quantitative psychologist or biostatistician who develops statistical learning methods for multivariate outcomes (e.g., educational test responses or psychiatric symptom scores) and creates pedagogical tools and scholarship to help applied researchers use these methods in practice.

• A scholar of languages or literature (modern or classical), who uses electronic corpora of texts for quantitative and qualitative analysis such as “Distant Reading” or corpus linguistics.

• A graph theory network analyst who can extract useable and actionable information from large scale pathway databases and match with experimental biological data.

• A scholar with deep evolutionary and computational biology knowledge who can integrate large-scale genomics/proteomics data across species (worm, fly, mouse, human gene expression).

• A systems biologist who uses machine learning to integrate functional genomic datasets across multiple scales.

• An engineering scholar (to not only do the work, but also to lead further initiatives) to develop big data, deep learning on medical imaging data to answer diagnostic as well as prognostic questions. The latter focused on surgical or therapeutic outcomes.

• A data scientist to develop algorithms on the patient record and harmonization of those records with large scale imaging, intervention, and outcome data.
Appendix H: Research IT Support for Data Science

An effective, usable, scalable, and sustainable IT infrastructure has been identified as an essential element for the success of a very broad array of research and scholarly activities at Vanderbilt. Discussions in the Data Science Visions Working Group have highlighted how Vanderbilt’s research IT infrastructure is particularly critical for advancing data science. Over the past year, the needs of the faculty at Vanderbilt have been identified and candidate solutions to infrastructure issues relevant to data science—and beyond—have been proposed by the Research IT Working Group, tasked in 2017 by the provost.

One recommendation to come out of the group was the establishment of an office that could supply a support structure, as well as the development of services to support faculty. As a result, the newly established Vanderbilt University Research IT Service (VU-RITS) has been charged with establishing research support by piloting three new services (described below). As described in this appendix, these VU-RITS services can help to address some of the “threats identified in the SWOT analysis section, including bandwidth, speed, storage, and general research IT.

Service 1: Sustainable Solution for Primary Research Data Storage and Backup
Data storage and backup is an area that needs immediate attention to achieve the goals of the data science vision. Typical data science research workflows include acquiring source data, deriving subsets of those data, manipulation of those data, sharing and storage of those data. All data need a safe, secure, reliable, affordable, and accessible location to reside within. As a solution to these workflows, both for data science and broader research needs on campus, the Research IT Working Group established a faculty focus group with vested interest in storage IT. These campus experts and stakeholders have evaluated the needs, assessed the current state-of-the-practice/art, and determined potential economies of scale.

Recommendations on research data storage, backup, and movement across key locations on campus through high-throughput research networks are currently being developed. Master agreements between Amazon Web Services (AWS) have been made with Vanderbilt, and workflows are being produced that ease connections between entities on campus, such as ACCRE and AWS, as well as other cloud providers, such as Microsoft Azure and Google Cloud.

Service 2: Research IT Consultation and Programming Services
The data science community, along with researchers and scholars across the broader campus community require technical assistance with project development and workflows, especially when workflows are data-intensive (as mentioned in the SWOT analysis section of this document). Faculty frequently have research accounts or grant funds but are often unwilling or unable to hire full-time staff. Such faculty would greatly benefit from access to shared technology experts, especially programmers with skills essential to conducting effective data science research and scholarship.

The Research IT Consultation and Programming Service aims to assist in quickly connecting researchers and scholars with experts (both identified by Research IT and housed within the
proposed Data Science Institute) that can aid in the development of custom programming solutions, especially with data-intensive workflows. As is the case with Pilot Service 2 discussed above, it is anticipated that the Vanderbilt Data Science Institute would be both a producer and consumer of the Research IT Consultation and Programming Service. In particular, this service aims to build on the data science visions effort, while also creating transferable approaches to research IT (by leveraging existing resources and those proposed by the Data Science Visions).

**Service 3: A Comprehensive Guide to Vanderbilt Research IT Solutions**

Research IT mechanisms and services have been implemented to help faculty (including, but certainly not limited to faculty interested in data science) by documenting research IT resources that are already available on campus. Work has been done to create and promote multiple avenues of connection between Vanderbilt scholars and the solutions to their identified or not-yet-identified IT needs. These avenues of connection include web-based guides that list campus resources, summarize what they provide, and offer contact information. The web-based guides also serve as a means through which both data science faculty and the broader research community may be connected to consultative, programming, and storage solutions identified by the Research IT working group.

The Vanderbilt Data Science Institute would be both a producer and consumer of research IT solutions available through the VU-RITS office, in general, and the web-based guide, in particular. For example, data scientists hired through the institute could be providers for the programmer for hire and consulting services outlined above.