**Program Announcement**
**New Jersey Institute of Technology**
**B.S. in Data Science**




Contact: **David A. Bader**, Distinguished Professor, Ying Wu College of Computing,
Director, Institute for Data Science, New Jersey Institute of Technology

## I. Objectives

New Jersey Institute of Technology (NJIT) proposes a Bachelor of Science in Data Science in order to prepare students for careers in the emerging and growing industry of Data Science, which has been disrupting every sector of industry through rapid integration of technology, algorithms, data and mobile applications. The program has two concentrations in Computing and Statistics.

Data science is the study and practice of extracting information and structure from data that can then be used for reasoning and adding value to the solution of a problem. It has growing applications in health and medicine, finance, marketing, economics, genomics, social networks, cyber-security, journalism, and other fields where data is collected. It spans academic fields in computer science and mathematics such as machine learning and statistical inference, probability, linear algebra, computer programming, software engineering, high performance computing, and cloud computing. In this region and nationally, there is a critical and growing need for producing a workforce skilled in data science throughout industry, labs, and government. Various faculty members in the Ying Wu College of Computing, the Department of Mathematical Sciences, and the broader university are already engaged in research that falls under the umbrella of data science. For example, in computer science there are faculty working in bioinformatics, medical informatics, image processing, data mining, big data, cyber-security, and cloud computing. In informatics there are faculty working in information integration, biomedical text mining, search engines, recommender systems, deep learning for ecological data, social media analysis, and knowledge engineering. Faculty from mathematical sciences, physics, management, and engineering are engaged in analyzing large datasets that usually involve statistical inference, machine learning, and high-performance computing. Thus, the scope of the field of data science extends across the university.

The B.S. in Data Science program with be co-managed by the Ying Wu College of Computing and the Department of Mathematical Sciences in the College of Science and Liberal Arts. Students will apply directly to the unit that offers the concentration (YWCC for Computing, and Mathematical Sciences for Statistics). The concentrations share significant coursework. Fourteen (14) courses are required by both concentrations, eleven (11) courses are required by one concentration and an elective in the other concentration, and nineteen (19) courses are shared electives. In addition, both concentrations require a two-semester senior project-based course. Students may request to transfer between the concentrations prior to their fourth semester.

## II. Evaluation & Assessment Plan for the Program

Currently, we are not aware of an appropriate accreditation program for data science undergraduate programs. We would seek accreditation from the Computing Accreditation Commission (CAC) of ABET should it assess B.S. in Data Science programs in the future. Nevertheless, the B.S. in Data Science will be assessed in accordance with the institution's existing assessment standards and practices to ensure continual program improvement.

The assessment is designed to measure students' foundational knowledge before graduation in their degree programs, and compares it against the learning goals identified. The data collected provide a detailed insight into the competencies and deficiencies of our students and programs. Assessment results are analyzed, disseminated, and utilized by the faculty for the purposes of curriculum planning and the improvement of teaching and learning.
The assessment plan involves the following steps:

- The identification and revision of learning goals, objectives and outcomes at the school level and for the Data Science degree program.
- The identification and development of direct and indirect assessment methods, instruments, and the criteria for proficiency.
- Annual data collection, analysis and reporting to the Undergraduate Curriculum Committee for curriculum planning and the improvement of teaching and learning.
- Closing the loop by integrating the recommended changes to the B.S. in Data Science curriculum.

Students graduating with a degree in Data Science should anticipate the acquisition of skills, knowledge, and professional training that enable them to pursue data science careers such as data scientist, data analyst, data engineer, in a broad range of industrial sectors and government. The primary goal of the B.S. in Data Science is to develop students who have the necessary skills and knowledge to pursue competitive professional and academic careers.

At the institutional level, the Office of Institutional Effectiveness (OIE) is responsible for assessment oversight at NJIT. The OIE works with individual academic divisions and units in order to assess academic programs on a regular basis in an effective, structured, and reliable manner.

The proposed B.S. in Data Science has been established based on program goals, learning objectives, as well as students learning outcomes. Figure 1 displays a summary of the assessment and evaluation plans. Tables 1 and 2 present the assessment plans for the program goals, objectives, and student learning outcomes, respectively, designed to achieve all related professional as well as institutional requirements.

### II.A. Data Science Program Goals

There are three comprehensive program goals (PGs).

1. Provide students with a broad systematic understanding of data science, as well as the application of this knowledge to improve the operational efficiency of organizations. [PG1]
2. Support NJIT's mission of excellence in education, research, economic development, and service. [PG2]

3. Develop and sustain a baccalaureate degree program that can broaden the existing computing disciplines, has growing and enduring professional demand, and is aligned with leading educational standards. [PG3]

## II.B. Data Science Program Objectives

The program comprises four student learning objectives (POs), which will be integrated into the existing assessment structure of the Ying Wu College of Computing (for the Computing concentration) and Department of Mathematical Sciences (for the Statistics concentration).

1. Understand foundational knowledge of the data science discipline. [PO1]
2. Demonstrate proficiency in using data science software and technology. [PO2]
3. Communicate effectively in a professional environment. [PO3]
4. Demonstrate critical/integrative thinking skills in data-driven problem-solving and decision making. [PO4]

## II.C. Data Science Student Learning Outcomes

In addition to goals and objectives, the proposed B.S. in Data Science includes two sets of student learning outcomes: (1) program-specific and (2) institutional (i.e., NJIT).

### II.C.1. Program-Based Student Learning Outcomes

The Data Science program-based learning outcomes (LOs) are organized by the program objectives articulated in Section II.B. Students earning a B.S. in Data Science from NJIT will demonstrate proficiency in their abilities to:

1.1. Describe key concepts in data science, including machine learning, artificial intelligence, big data, and analytics. [LO1]
1.2. Construct data models and apply them in product design, strategic analyses, and improving operational efficiency of organizations. [LO2]
1.3. Discuss ethics and regulations as they relate to Data Science, such as privacy and security, legitimate use of data, both nationally and internationally. [LO3]

2.1. Demonstrate advanced computing skills for data-driven decision making. [LO4]
2.2. Demonstrate effective programming skills using languages such as Python, R, and C to perform data analysis. [LO5]
2.3. Apply technologies proficiently offered for analytical software, modeling software, and machine learning tools. [LO6]
2.4. Use document collaboration tools proficiently, such as Google Docs, OneDrive, Dropbox, and communication tools such as email, social media, and Zoom/Webex, and career development tools such as LinkedIn. [LO7]

3.1. Write organized, clear and concise reports, addressing concepts, and analyses. [LO8]
3.2. Make effective, clear, and informative presentations using data science tools. [LO9]

4.1. Consider facts and identify problems accurately. [LO10]
4.2. Evaluate alternative solutions to problems and propose possible actions. [LO11]
4.3. Implement and evaluate large-scale real-world projects. [LO12]
4.4. Demonstrate the ability to work in a team. [LO13]

For a summary of the program-based student learning outcomes, and in order to view their mapping onto the program objectives, please see Tables 2 and 3.

II.C.2. NJIT's General Education Requirement (GER) Student Learning Outcomes

NJIT is dedicated to producing graduates who have the knowledge, skills, and motivation necessary to advance the state-of-the-art knowledge in their respective fields in addition to possessing a devotion to lifelong personal development as well as intellectual discovery beyond their discipline. Graduates must possess outstanding communication skills and understand the complexities of contemporary society and the ethical and societal issues involved in the professional pursuit of their discipline. Graduates must also possess a deep understanding of and appreciation for science and technology. The General Education Requirements (GER) are designed to be the dynamic yet minimal foundational curriculum encompassing the necessary preconditions for success in undergraduate disciplines as well as the breadth of knowledge demanded by contemporary society.

There are five NJIT GER student learning outcomes (GERs):

1. Effectively communicate ideas orally and in writing, as informed by the tenets of a liberal arts education (Liberal Arts Literacy). [GER1]
2. Use logical reasoning and a scientific approach to support conclusions based on empirical evidence (Scientific Literacy). [GER2]
3. Form conclusions that are supported logically by the principles of qualitative and quantitative reasoning, probability, and statistics (Quantitative Literacy). [GER3]
4. Demonstrate the ability to use computing systems in order to access, store, process and analyze information as an essential aspect of critical thinking and problem solving (Computing Literacy). [GER4]
5. Identify and articulate the multifaceted relationships between the economic, social and political forces that inform and structure society as well as an individual's place within it (Social Science Literacy). [GER5]

For a summary of the institutional student learning outcomes, and in order to view their mapping onto the program objectives, please see Table 2 (for the Computing concentration) and Table 3 (for the Statistics concentration).

***Figure 1. Program Evaluation and Outcome Assessment Hierarchical Summary***

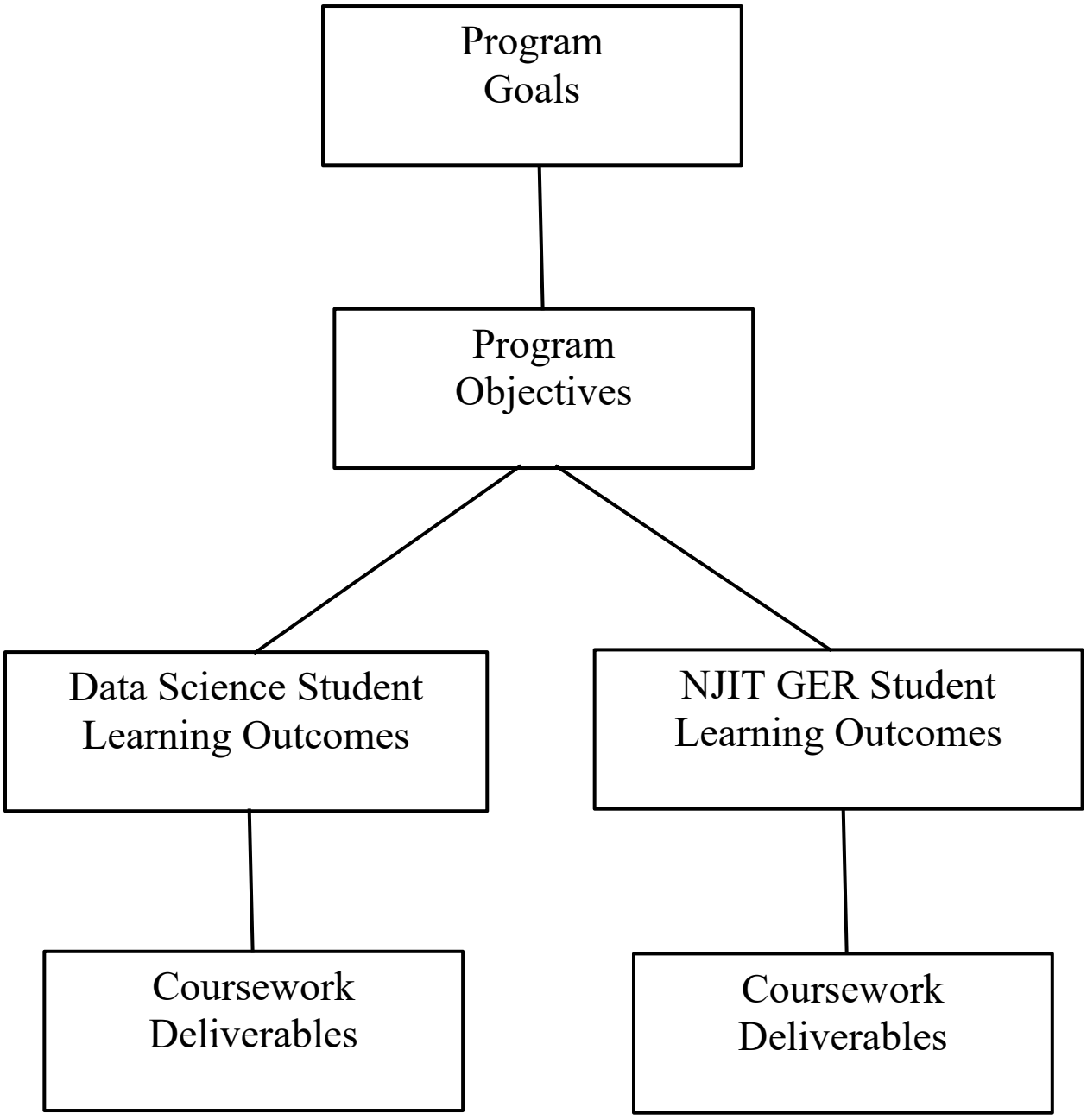

**Table 1. Evaluation of Data Science Program Goals**

| Program Goals | Measures |
|---|---|
| **PG1:** Provide students with a broad systematic understanding of data science, as well as the application of this knowledge to improve the operational efficiency of organizations. | ▪ Enrollment and retention data.<br>▪ Student satisfaction surveys.<br>▪ Student, alumni, and employer surveys ascertaining employment, career placement, and career advancement.<br>▪ Exit interviews.<br>▪ Degree completion data. |
| **PG2:** Support NJIT's mission of excellence in education, research, economic development, and service. | ▪ Appropriate institutional and state-level (i.e., AIC/NJPC) approvals.<br>▪ Successful recruitment and enrollment of gender and racially diverse students. |
| **PG3:** Develop and sustain a baccalaureate degree program that can broaden the existing computing disciplines, has growing and enduring professional demand, and is aligned with leading educational standards. | ▪ Surveys ascertaining employment and salary 1, 5, and 10 years after graduating the program.<br>▪ Regularly scheduled program assessment and review. |

## Table 2. Curriculum Map & Assessment of Programmatic & Institutional Student Learning Outcomes (Computing concentration)

| Program Objectives | Student Learning Outcomes | Courses | Assessment |
|---|---|---|---|
| **PO1:** Understand foundational knowledge of the data science discipline. | **LO1:** Understand key concepts in data science, including machine learning, artificial intelligence, big data, and analytics.<br>**LO2:** Understand data models and their use in product design, strategic analyses, and improving operational efficiency of organizations.<br>**LO3:** Understand ethics and regulations around Data Science, such as privacy and security, legitimate use of data, both nationally and internationally.<br>**GER3:** Form conclusions that are supported logically by the principles of qualitative and quantitative reasoning, probability, and statistics.<br>**GER5:** Identify and articulate the multifaceted relationships between the economic, social and political forces that inform and structure society as well as an individual's place within it. | CS 100, CS 113, CS 114, CS 241, CS 280, CS 301, CS 351, IS 350, MATH 111, MATH 112, MATH 244, MATH 337, MATH 344, MATH 478 | Objective exam results<br>Projects with rubrics<br>Case studies with rubrics<br>Homework assignments<br>Quizzes with rubrics<br>Lab reports with rubrics |
| **PO2:** Demonstrate proficiency in using data science software and technology. | **LO4:** Demonstrate advanced computing skills for data-driven decision making.<br>**LO5:** Demonstrate effective programming skills using languages such as Python, R and C to perform data analysis.<br>**LO6:** Apply technologies proficiently offered for analytical software, modeling software, and machine learning tools.<br>**LO7:** Use document collaboration tools proficiently, such as Google Docs, OneDrive, Dropbox, and communication tools such as email, social media, and Zoom/Webex, and career development tools such as LinkedIn.<br>**GER4:** Demonstrate the ability to use computing systems in order to access, store, process and analyze information as an essential aspect of critical thinking and problem solving (Computing Literacy). | CS 288, CS 331, CS 370, CS 375 (Machine Learning), CS 450 (Data Visualization), CS 444 (Big Data Systems), CS 435, CS 482. MATH 341 | Lab reports with rubrics<br>Projects with rubrics<br>Objective exam results<br>Homework assignments<br>Case studies with rubrics<br>Quizzes with rubrics |
| **PO3:** Communicate effectively in a professional environment. | **LO8:** Write organized, clear and concise reports, addressing concepts, and analyses.<br>**LO9:** Make effective, clear, and informative presentations using financial technology tools.<br>**GER1:** Effectively communicate ideas orally and in writing, as informed by the tenets of a liberal arts education. | YWCC 207, ENG 340/352, HUM 101, HUM 102, MGMT 216 | Oral and written assignments with rubrics<br>One-on-one training with rubrics<br>Case studies with rubrics<br>Objective exam results |
| **PO4:** Demonstrate critical/integrative thinking skills in data-driven problem-solving and decision making. | **LO10:** Consider facts and identify problems accurately.<br>**LO11:** Evaluate alternative solutions to problems and propose possible actions.<br>**LO12:** Implement and evaluate large-scale real-world projects.<br>**LO13:** Demonstrate the ability to work in a team.<br>**GER2:** Use logical reasoning and a scientific approach to support conclusions based on empirical evidence (Scientific Literacy). | YWCC 307, CS 492 (Data Science Capstone 1), CS 493 (Data Science Capstone 2) | Quizzes with rubrics<br>Case studies with rubrics<br>Homework assignments<br>Objective exam results<br>Lab reports with rubrics<br>Semester long collaborative project |

## Table 3. Curriculum Map & Assessment of Programmatic & Institutional Student Learning Outcomes (Statistics concentration)

| Program Objectives | Student Learning Outcomes | Courses | Assessment |
|---|---|---|---|
| **PO1:** Understand foundational knowledge of the data science discipline. | **LO1:** Understand key concepts in data science, including machine learning, artificial intelligence, big data, and analytics.<br>**LO2:** Understand data models and their use in product design, strategic analyses, and improving operational efficiency of organizations.<br>**LO3:** Understand ethics and regulations around Data Science, such as privacy and security, legitimate use of data, both nationally and internationally.<br>**GER3:** Form conclusions that are supported logically by the principles of qualitative and quantitative reasoning, probability, and statistics.<br>**GER5:** Identify and articulate the multifaceted relationships between the economic, social and political forces that inform and structure society as well as an individual's place within it. | CS 100, CS 113, CS 114, CS 241, CS 280, CS 301, MATH 111, MATH 112, MATH 213, MATH 244, MATH 337, MATH 344, MATH 345, MATH 391, MATH 447, MATH 448, MATH 478 | Objective exam results<br>Projects with rubrics<br>Case studies with rubrics<br>Homework assignments<br>Quizzes with rubrics<br>Lab reports with rubrics |
| **PO2:** Demonstrate proficiency in using data science software and technology. | **LO4:** Demonstrate advanced computing skills for data-driven decision making.<br>**LO5:** Demonstrate effective programming skills using languages such as Python, R and C to perform data analysis.<br>**LO6:** Apply technologies proficiently offered for analytical software, modeling software, and machine learning tools.<br>**LO7:** Use document collaboration tools proficiently, such as Google Docs, OneDrive, Dropbox, and communication tools such as email, social media, and Zoom/Webex, and career development tools such as LinkedIn.<br>**GER4:** Demonstrate the ability to use computing systems in order to access, store, process and analyze information as an essential aspect of critical thinking and problem solving (Computing Literacy). | CS 450 (Data Visualization), MATH 340, MATH 341, MATH 461 | Lab reports with rubrics<br>Projects with rubrics<br>Objective exam results<br>Homework assignments<br>Case studies with rubrics<br>Quizzes with rubrics |
| **PO3:** Communicate effectively in a professional environment. | **LO8:** Write organized, clear and concise reports, addressing concepts, and analyses.<br>**LO9:** Make effective, clear, and informative presentations using financial technology tools.<br>**GER1:** Effectively communicate ideas orally and in writing, as informed by the tenets of a liberal arts education. | ENG 340/352, HUM 101, HUM 102, MGMT 216 | Oral and written assignments with rubrics<br>One-on-one training with rubrics<br>Case studies with rubrics<br>Objective exam results |
| **PO4:** Demonstrate critical/integrative thinking skills in data-driven problem-solving and decision making. | **LO10:** Consider facts and identify problems accurately.<br>**LO11:** Evaluate alternative solutions to problems and propose possible actions.<br>**LO12:** Implement and evaluate large-scale real-world projects.<br>**LO13:** Demonstrate the ability to work in a team.<br>**GER2:** Use logical reasoning and a scientific approach to support conclusions based on empirical evidence (Scientific Literacy). | MATH 462 (Statistics Capstone 1), MATH 463 (Statistics Capstone 2) | Quizzes with rubrics<br>Case studies with rubrics<br>Homework assignments<br>Objective exam results<br>Lab reports with rubrics |

### II.D. Institutional Learning Goals

Demonstrate critical/integrative thinking skills in data-driven problem-solving and decision making.

In compliance with the accreditation standards and guidelines of the Middle States Commission on Higher Education, NJIT maintains adherence to the following five Institutional Learning Goals (ILGs):

1. Research-Based Inquiry: Students employ methods appropriate to their discipline. [ILG1]
2. Collaboration: Students work effectively in teams, applying multidisciplinary and global perspectives. [ILG2]
3. Ethical Conduct: Students demonstrate professional and civic responsibility, including respect for all individuals. [ILG3]
4. Creativity: Students use heuristics to evaluate, analyze, and synthesize innovative solutions to existing and emerging problems. [ILG4]
5. Professional Readiness: Students exhibit knowledge and skills, and engage in experiences, necessary for professional and personal growth. [ILG5]

The assessment is handled by dedicated faculty and committees. The accreditation committee, the undergraduate curriculum committee, the academic advisor, and the Associate Dean of Academic affairs will work collaboratively to collect and analyze data, as well as develop guidelines and implement changes to improve student learning, and thus closing the loop.

## III. Relationship of the Program to the Institutional Strategic Plan & Institutional Impact

### III.A. Relationship to Institutional Strategic Plan

NJIT is one of the elite polytechnic universities in New Jersey that is dedicated to engineering, applied science and the management of technology in education. The university programs are career-focused and conduct applied learning. Its strategic plan is based on five strategic priorities, one of which is learning. The learning objectives are:

- *To ensure it[course of study] meets current professional standards, provides a general education to produce the most highly qualified leaders and is delivered by enthusiastic instructors using innovative and effective methods.*
- *To ensure degree programs meet the needs of students and the employment market.*
- *Train students in the competencies of current digital technology related to their majors and integrate this into the general university requirements. This will provide graduates with technological knowledge and skills required by the marketplace.*

Clearly, the proposed B.S. in Data Science degree program aligns well with the objectives of NJIT's technology focused education, and has the same goal of providing students the skills required by multiple industries, and will contribute to New Jersey's economic and workforce development.

### III.B. Impact on Existing Program

NJIT encourages colleges and programs to develop new programs to maximize the availability of expertise to students in different fields.

The B.S. in Data Science program will be managed by the Ying Wu College of Computing, and coordinated with the Department of Mathematical Sciences. The program will leverage existing coursework, and enhance the data science offerings at NJIT. Other programs such as B.S. in Computer Science, and B.S. in Financial Technology in MTSM, may benefit from these new course offerings.

## IV. Need

Data Science is critical to national competitiveness, and as such, NJIT has launched the Institute for Data Science. In the recent National Science Foundation report (REALIZING THE POTENTIAL OF DATA SCIENCE Final Report from the National Science Foundation Computer and Information Science and Engineering Advisory Committee Data Science Working Group. Francine Berman and Rob Rutenbar, co-Chairs Henrik Christensen, Susan Davidson, Deborah Estrin, Michael Franklin, Brent Hailpern, Margaret Martonosi, Padma Raghavan, Victoria Stodden, Alex Szalay. December 2016), it is clear: "The ability to manipulate data and understand Data Science is becoming increasingly critical to current and future discovery and innovation." Further, McKinsey predicts that data-driven technologies will bring an additional $300 billion of value to the U.S. health care sector alone, and by 2020, 1.5 million more "data-savvy managers" will be needed to capitalize on the potential of data, "big" and otherwise. According to Business Insider, data scientist continues to rank as the best jobs in America, with an average compensation of $120,000 per year (Glassdoor).

The ACM Education Board launched the ACM Data Science Task Force to develop an undergraduate curriculum in Data Science. Education Board co-chairs Chris Stephenson and Jane Prey wrote in their charge to the task force: "In 2009, Turing award winner Jim Gray spoke of data science as a fourth paradigm of science (empirical, theoretical, computational and data-driven) arising from and capitalizing on the huge amount of data that is now available for investigation. The confluence of the availability of data and increasing sophisticated tools, processes, and algorithms for analyzing and drawing knowledge and insight from data has impacted every area of scientific engagement. It has also opened up exciting new opportunities for interdisciplinary work across the many fields including (but certainly not limited to) computer science, mathematics, statistics, and information science from which it draws foundational knowledge."

The ACM Data Science Task Force, recognizing the high demand for a skilled workforce in data science, produced a draft report (version 2, December 2019) on "Computing Competencies for Undergraduate Data Science Curricula" that informs the curriculum in this proposal. http://dstf.acm.org/

Experts recognize the growing demand for undergraduate students trained in data science:

- PwC: What's next for the data science and analytics job market?
  https://www.pwc.com/us/en/library/data-science-and-analytics.html
  US Employers want 86% of analytics-enabled jobs and 59% of data scientist and advanced analyst jobs to have a bachelors degree.
- Irving Wladawsky-Berger: Why Do We Need Data Science when We've Had Statistics for Centuries?
  https://blog.irvingwb.com/blog/2014/04/why-do-we-need-data-science-when-weve-had-statistics-for-centuries.html
- Forbes: LinkedIn's Fastest-Growing Jobs Today Are In Data Science And Machine Learning
  https://www.forbes.com/sites/louiscolumbus/2017/12/11/linkedins-fastest-growing-jobs-today-are-in-data-science-machine-learning/#794e545551bd
  "Job growth in the next decade is expected to outstrip growth during the previous decade, creating 11.5M jobs by 2026, according to the U.S. Bureau of Labor Statistics."
- Wall Street Journal, 30 May 2020, AI Isn't Magical and Won't Help You Reopen your Business
  https://www.wsj.com/articles/ai-isnt-magical-and-wont-help-you-reopen-your-business-11590811201
  "Google CEO Sundar Pichai has said that in the sweep of human history, AI is more important than electricity or fire, and all the Big Five have said they'll continue to add to their engineering ranks during the downturn, including data scientists and AI experts."
- Wall Street Journal, 30 March 2020: AI Hiring Expected to Show Resilience Amid Coronavirus Slowdown
  https://www.wsj.com/articles/ai-hiring-expected-to-show-resilience-amid-coronavirus-slowdown-11585609823
  "Technology workers with artificial-intelligence skills are being spared from the current wave of layoffs and may see employment opportunities increase"

NJIT is located in the heart of the region where data scientist jobs are most needed in the United States. According to the U.S. Bureau of Labor Statistics, May 2019 Occupational Employment and Wages Report, the New York-Newark-Jersey City, NY-NJ-PA, metropolitan region has the highest number of jobs for Data Scientists and Mathematical Scientists with 2,690 employees, and an annual mean wage of $125,030.
https://www.bls.gov/oes/current/oes152098.htm

## IV.A. Impact on Regional Economy

The State of New Jersey has been investing in workforce education and school systems, to encourage financial know-how. Growing the state's Data Science Ecosystem has been an interest of the Governor and the New Jersey Economic Development Authority. The proposed degree will increase student enrollment from across the state, internationally, and around the Greater New York area. It will also generate a skilled workforce and attract Data Science startups to the region.

## IV.B. Relationship to Other Programs in the State & Region

Only a few Data Science programs at the levels of B.S. and M.S. are currently offered in the State:

- Ramapo College of New Jersey offers a BS in Data Science program: https://www.ramapo.edu/data-science/bachelor-of-science/
- NJCU offers a BS in Data Science program: https://www.njcu.edu/academics/schools-colleges/school-business/departments/finance/undergraduate-programs/business-analytics-and-data-science-bs
- Rutgers University offers an MS in Data Science degree program (within the departments of Computer Science and Statistics & Biostatistics in New Brunswick), but no undergraduate degree.
  https://www.rutgers.edu/academics/explore-undergraduate-programs
  https://grad.admissions.rutgers.edu/GraduateProgram/
- Stevens Institute of Technology offers an MS in Data Science through the Schaefer School of Engineering & Science, but does not offer any undergraduate data science degree programs.
  https://www.stevens.edu/academics/undergraduate-studies/undergraduate-majors
  https://www.stevens.edu/academics/graduate-education/graduate-programs
- Princeton University does not offer any data science degree programs.
  https://admission.princeton.edu/academics/degrees-departments
  https://gradschool.princeton.edu/academics/fields-of-study

Nationally, other schools with B.S. in Data Science programs include: UC San Diego, Tufts University, Indiana University, Temple University, DePaul University, Drexel University, Northeastern University, University of Michigan, UC Berkeley, Purdue University, Pennsylvania State University, Iowa State University, and Columbia University.

Other schools with Data Science programs are listed at:

https://www.discoverdatascience.org/programs/bachelors-in-data-science/

## V. Student Enrollment

The prospective students who would be enrolled in the proposed program are high school graduates who are interested in data science, county college students in associate degree programs in computer science or data science certificates, non-degree students who minor in computing or data science, and computing college students who have the programming and computer science background and are interested in a career in data science.

Projections for the proposed B.S. in Data Science program (see Table 4) are based on anticipated demand. With program marketing to capitalize on innate interest in the degree program, we estimate that the program will have at least 80 majors by the end of the second academic year that it is offered. Building on this momentum, we project an approximate additional 40 majors in the third year, an approximate additional 50 majors in the fourth year, and an approximate additional 60 majors in the program's fifth year. Please note that in year one we anticipate that the 20 students currently pursuing the B.S. degrees in Computer Science and Information Technology in the Ying Wu College of Computing, and Mathematical Sciences in the College of Science and Liberal Arts, will opt for the new major and are thus counted as continuing students.

**Table 4. B.S. in Data Science Enrollment Projections, AYs 2021-2026**

| Academic year | New Freshmen Enrollment | Transfer Student Enrollment | Continuing Student Enrollment |
|---|---|---|---|
| 2021-2022 | **20** | **0** | **20** |
| 2022-2023 | **30** | **10** | **40** |
| 2023-2024 | **40** | **15** | **80** |
| 2024-2025 | **50** | **15** | **135** |
| 2025-2026 | **60** | **15** | **200** |

### V.A. Admission Requirements

There are no special requirements for admission to the B.S. in Data Science program. As with our other undergraduate programs in the YWCC and CSLA and in complying with NJIT's academic policies, any matriculated student may declare Data Science as a major. Students must maintain a "C" average in their upper-division major courses in order to be certified for graduation. As described previously, students will apply to the unit that serves as the concentration home (YWCC for the Computing concentration and Mathematical Sciences for Statistics concentration). Prior to the fourth semester, students may request to transfer between the concentrations.

## VI. Program Resources

In order for the proposed B.S. in Data Science to be a successful program, it is important that the necessary institutional resources are identified and allocated properly. To this end, resources across the following areas will be considered in this section: coursework, faculty, libraries and computing facilities, and classrooms and laboratories.

### VI.A. Coursework & Course Development

NJIT is in a strategic position to launch an undergraduate degree program in Data Science. The following courses that are required by the curriculum of the proposed program are currently being offered in the Ying Wu College of Computing and Department of Mathematical Sciences at NJIT.

Curriculum Structure for BS in Data Science (Computing Concentration)

1. Required Courses:
    a. Foundational: (5 courses)
        1. CS 100 (Roadmap to Computing)
        2. MATH 111 (Calculus I)
        3. MATH 112 (Calculus II)
        4. MATH 244 (Intro to Probability Theory)
        5. MATH 341 (Statistical Methods)
    b. CS Core: (13 courses)
        i. Computing and Computer Fundamentals
            1. CS 113 (Introduction to Computer Science)
            2. CS 114 (Introduction to Computer Science II)
            3. CS 241 (Foundations of Computer Science I)
        ii. Programming, Data Structures, and Algorithms
            1. CS 280 (Programming Language Concepts)
            2. CS 288 (Intensive Programming in Linux)
            3. CS 435 (Advanced Data Structures and Algorithm Design)
        iii. Data Analysis, Presentation, Management, and Systems
            1. CS 301 (Introduction to Data Science)
            2. CS 331 (Database System Design & Management)
            3. CS 370 (Introduction to Artificial Intelligence)
            4. CS 482 (Data Mining)
            5. CS 375 (Machine Learning)
            6. CS 450 (Data Visualization)
            7. CS 444 (Big Data Systems)
    c. MATH Core: (3 courses)
        1. MATH 337 (Linear Algebra)
        2. MATH 344 (Regression Analysis)
        3. MATH 478 (Stat Methods in Data Science)

d. Applications and Software Development of Data Science Core: (2 courses)
    1. CS 492 (Data Science Capstone 1)
    2. CS 493 (Data Science Capstone 2)

2. Electives: (4 Courses)
    1. CS 310 (Co-op Work Experience - Data Science Position)
    2. CS 332 (Principles of Operating Systems)
    3. CS 350 (Introduction to Computer Systems)
    4. CS 351 (Introduction to Cybersecurity)
    5. CS 356 (Introduction to Computer Networks)
    6. CS 357 (Fundamentals of Network Security)
    7. CS 408 (Cryptography and Internet Security)
    8. IS 333 (Social Network Analysis)
    9. IS 392 (Web Mining and Information Retrieval)
    10. IT 430 (Ethical Hacking for System Administrators)
    11. IT 485 (ST: Deep Learning in Games)
    12. MATH 345 (Multivariate Distributions)
    13. MATH 388 (Intro to Chaos Theory)
    14. MATH 391 (Numerical Linear Algebra)
    15. MATH 430 (Analytical & Computational Neuroscience)
    16. MATH 447 (Applied Time Series Analysis)
    17. MATH 448 (Stochastic Simulation)
    18. MATH 461 (Statistical Computing with R)
    19. MGMT 216 (Business Data Analytics)
    20. MGMT 316 (Business Research Methods)
    21. MGMT 416 (Artificial Intelligence for Business Decisions)
    22. MRKT 378 (Marketing Analytics)
    23. MRKT 430 (Marketing Research)
    24. FIN 218 (Financial Markets and Institutions)
    25. FIN 306 (Blockchain Technology for Business)
    26. FIN 310 (Data Driven Financial Modeling)
    27. FIN 320 (Fin Data Analytics with R Prog)

Curriculum Structure for BS in Data Science (Statistics Concentration)

1. Required Courses:
    a. Foundational: (5 courses)
        1. CS 100 (Roadmap to Computing)
        2. MATH 111 (Calculus I)
        3. MATH 112 (Calculus II)
        4. MATH 244 (Intro to Probability Theory)
        5. MATH 341 (Statistical Methods)
    a. CS Core: (6 courses)

i. Computing and Computer Fundamentals
    1. CS 113 (Introduction to Computer Science)
    2. CS 114 (Introduction to Computer Science II)
    3. CS 241 (Foundations of Computer Science I)
ii. Programming, Data Structures, and Algorithms
    1. CS 280 (Programming Language Concepts)
iii. Data Analysis, Presentation, Management, and Systems
    1. CS 301 (Introduction to Data Science)
    2. CS 450 (Data Visualization)

b. MATH Core: (10 courses)
    1. MATH 213 (Calculus IIIB)
    2. MATH 337 (Linear Algebra)
    3. MATH 340 (Applied Numerical Methods)
    4. MATH 344 (Regression Analysis)
    5. MATH 345 (Multivariate Distributions)
    6. MATH 391 (Numerical Linear Algebra)
    7. MATH 447 (Applied Time Series Analysis)
    8. MATH 448 (Stochastic Simulation)
    9. MATH 461 (Statistical Computing with R)
    10. MATH 478 (Stat Methods in Data Science)

c. Applications and Software Development of Data Science Core: (2 courses)
    1. MATH 462 (Statistics and Statistical Learning Capstone 1)
    2. MATH 463 (Statistics and Statistical Learning Capstone 2)

2. Electives: (4 Courses)
    1. CS 310 (Co-op Work Experience - Data Science Position)
    2. CS 331 (Database System Design and Management)
    3. CS 332 (Principles of Operating Systems)
    4. CS 350 (Introduction to Computer Systems)
    5. CS 351 (Introduction to Cybersecurity)
    6. CS 356 (Introduction to Computer Networks)
    7. CS 357 (Fundamentals of Network Security)
    8. CS 370 (Introduction to Artificial Intelligence)
    9. CS 408 (Cryptography and Internet Security)
    10. CS 435 (Advanced Data Structures and Algorithm Design)
    11. CS 482 (Data Mining)
    12. CS 450 (Machine Learning)
    13. CS 444 (Big Data Systems)
    14. IS 333 (Social Network Analysis)
    15. IS 392 (Web Mining and Information Retrieval)
    16. IT 430 (Ethical Hacking for System Administrators)
    17. IT 485 (ST: Deep Learning in Games)
    18. MATH 222 (Intro to Differential Equations)
    19. MATH 388 (Intro to Chaos Theory)

20. MATH 430 (Analytical & Computational Neuroscience)
21. MATH 453 (High-Performance Numerical Computing)
22. MATH 477 (Stochastic Processes)
23. MGMT 316 (Business Research Methods)
24. MGMT 416 (Artificial Intelligence for Business Decisions)
25. MRKT 378 (Marketing Analytics)
26. MRKT 430 (Marketing Research)
27. FIN 218 (Financial Markets and Institutions)
28. FIN 306 (Blockchain Technology for Business)
29. FIN 310 (Data Driven Financial Modeling)
30. FIN 320 (Fin Data Analytics with R Prog)

Curriculum of BS in Data Science (Computing concentration) (120 credits)

| **1st year (Fall) (14 credits)** | **1st year (Spring) (14 credits)** |
|---|---|
| CS 100: Roadmap to Computing | CS 113: Intro to Computer Science I |
| Math 111: Calculus I | Math 112: Calculus II |
| GER: Scientific Literacy with Lab 1 | GER: Scientific Literacy with Lab 2 |
| Hum 101: English Composition | Hum 102: English Composition |
| Freshman Seminar | |
| | |
| **2nd year (Fall) (15 credits)** | **2nd year (Spring) (16 credits)** |
| CS 114: Intro to Computer Science II | CS 241: Foundations of Computer Science I |
| Math 244: Introduction to Probability | CS 280: Programming Language Concepts |
| Math 337: Linear Algebra | IS 350: Computers, Society, and Ethics |
| GER: Humanities/History (200 level) | Math 341: Statistical Methods |
| GER: Social Science Literacy GER | YWCC 207: Computing & Effective Communication |
| | Data Science Elective 1 [3] |
| | |
| **3rd year (Fall) (15 credits)** | **3rd year (Spring) (16 credits)** |
| CS 288: Intensive Programming in Linux | CS 435: Advanced Data Structures and Algorithm Design |
| CS 301: Introduction to Data Science | Data Science Elective 2 [3] |
| CS 331: Database System Design & Management | CS 482; Data Mining |
| CS 370: Intro to Artificial Intelligence | CS 375: Machine Learning |
| ENG 340 or ENG 352: Oral Presentation or Technical Writing | GER: History and Humanities (300+ level) |
| | YWCC 307: Professional Development in Computing |
| | |
| **4th year (Fall) (15 credits)** | **4th year (Spring) (15 credits)** |
| CS 450: Data Visualization | GER: Humanities/Social Science Senior Seminar |
| CS 444: Big Data Systems | Free Elective 1 [2] |
| CS 492: Data Science Capstone 1 | CS 493: Data Science Capstone 2 |
| MATH 478: Stats Methods in Data Science | Math 344: Regression Analysis |
| Data Science Elective 3 [3] | Data Science Elective 4 [3] |
| | |
| | |

Curriculum of BS in Data Science (Statistics concentration) (120 credits)

| **1st year (Fall) (14 credits)** | **1st year (Spring) (14 credits)** |
|---|---|
| CS 100: Roadmap to Computing | CS 113: Intro to Computer Science I |
| Math 111: Calculus I | Math 112: Calculus II |
| GER: Scientific Literacy with Lab 1 [1] | GER: Scientific Literacy with Lab 2 [1] |
| Hum 101: English Composition | Hum 102: English Composition |
| Freshman Seminar | |
| | |
| **2nd year (Fall) (15 credits)** | **2nd year (Spring) (16 credits)** |
| CS 114: Intro to Computer Science II | CS 241: Foundations of Computer Science I |
| Math 244: Introduction to Probability | CS 280: Programming Language Concepts |
| Math 337: Linear Algebra | Math 213: Calculus IIIB |
| GER: Humanities/History (200 level) | Math 341: Statistical Methods |
| GER: Social Science | Data Science Elective 1 [4] |
| | |
| **3rd year (Fall) (15 credits)** | **3rd year (Spring) (15 credits)** |
| Math 340: Applied Numerical Methods | Math 345: Multivariate Distributions |
| Math 344: Regression Analysis | Math 447: Applied Time Series Analysis |
| Math 391: Numerical Linear Algebra | Math 478: Statistical Methods in Data Science |
| CS 301: Introduction to Data Science | Data Science Elective 2 [4] |
| GER: History and Humanities (300 level) | GER: History and Humanities (300 level) |
| | |
| **4th year (Fall) (15 credits)** | **4th year (Spring) (16 credits)** |
| Math 448: Stochastic Simulation | GER: Humanities/Social Science Senior Seminar |
| Math 461: Statistical Computing with R | Math Upper Level Elective (300+ level) |
| Math 462: Statistics and Statistical Learning Capstone 1 | Math 463: Statistics and Statistical Learning Capstone 2 |
| CS 450: Data Visualization | Data Science Elective 4 [4] |
| Data Science Elective 3 [4] | Free Elective (4 Credits) [2] |
| | |

**Notes**

1. Students considering switching to Computer Science or Mathematical Sciences should take PHYS 111/111A and 121/121A. Do not take PHYS 102/102A
2. Free electives should be chosen in consultation with the advisor. Some restrictions apply.
3. Choose from the following for the Data Science – Computing concentration (courses used to satisfy another requirement of the degree may not be counted as a Data Science Elective):

| | |
|---|---|
| CS 310: Co-op Work Experience | MATH 345: Multivariate Distributions |

| | |
|---|---|
| CS 332: Principles of Operating Systems | MATH 388: Intro to Chaos Theory |
| CS 350: Intro to Computer Systems | MATH 391: Numerical Linear Algebra |
| CS 351: Intro to Cybersecurity | MATH 430: Analytical/Comp Neuroscience |
| CS 356: Intro to Computer Networks | MATH 447: Applied Time Series Analysis |
| CS 357: Fundamentals of Network Security | MATH 448: Stochastic Simulation |
| CS 370: Intro to Artificial Intelligence | MATH 461: Statistical Computing with R |
| CS 375: Machine Learning | IS 333: Social Network Analysis |
| CS 444: Big Data Systems | IS 392: Web Mining & Information Retrieval |
| CS 408: Cryptography and Internet Security | FIN 218: Financial Markets and Institutions |
| | FIN 310: Data Driven Financial Modeling |
| MGMT 316: Business Research Methods | FIN 320: Financial Data Analysis with R |
| MRKT 378: Marketing Analytics | IT 430: Ethical Hacking for System Administrators |
| MRKT 430: Marketing Research | IT 485: ST Deep Learning in Games |
| FIN 306: Blockchain Technology for Business | MGMT 416: Artificial Intelligence for Business Decisions |

4. Choose from the following for the Data Science – Statistics concentration (courses used to satisfy another requirement of the degree may not be counted as a Data Science Elective):

| | |
|---|---|
| CS 310: Co-op Work Experience | MATH 222: Intro to Differential Equations |
| CS 331: Database System Design and Management | MATH 388: Intro to Chaos Theory |
| CS 332: Principles of Operating Systems | MATH 430: Analytical/Comp Neuroscience |
| CS 350: Intro to Computer Systems | MATH 453: High-Performance Numerical Computing |
| CS 351: Intro to Cybersecurity | MATH 477: Stochastic Processes |
| CS 356: Intro to Computer Networks | IS 333: Social Network Analysis |
| CS 357: Fundamentals of Network Security | IS 392: Web Mining & Information Retrieval |
| CS 370: Intro to Artificial Intelligence | FIN 218: Financial Markets and Institutions |
| CS 375: Machine Learning | FIN 310: Data Driven Financial Modeling |
| CS 444: Big Data Systems | FIN 320: Financial Data Analysis with R |
| CS 408: Cryptography and Internet Security | IT 430: Ethical Hacking for System Administrators |
| CS 435: Advanced Data Structures and Algorithm Design | IT 485: ST Deep Learning in Games |
| CS 482: Data Mining | |
| | FIN 306: Blockchain Technology for Business |
| MGMT 316: Business Research Methods | MGMT 416: Artificial Intelligence for Business Decisions |
| MRKT 378: Marketing Analytics | |
| MRKT 430: Marketing Research | |

## VI.B. Faculty

The most important aspect of offering a new degree in Data Science is the qualifications of the faculty who will be teaching for the program. The Data Science area requires an interdisciplinary background in not only traditional computer science concepts, but also in data modeling, data analysis, applied statistics, and programming. In the last five years, NJIT has invested in faculty recruitment from new and cutting-edge research areas, including artificial intelligence, machine learning, and big data. The following faculty will be participating in the Data Science program.

**Table 5. A List of Full-time Faculty Who Will Participate in the Proposed Program**

| Last Name | First Name | Title | Department | Research Areas |
|---|---|---|---|---|
| Bader | David A. | Distinguished Professor | Computer Science | High performance computing, computational biology, massive-scale data analytics, graph algorithm |
| Bandera | Cesar | Associate Professor of Entrepreneurship | MTSM | entrepreneurship education, innovation-driven business incubation, and mobile healthcare innovation |
| Basu Roy | Senjuti | Assistant Professor | Computer Science | Big Data analytics, data mining, machine learning, emerging applications |
| Borcea | Cristian | Professor | Computer Science | Cloud Computing, networks |
| Buntain | Cody | Assistant Professor | Informatics | Crisis Informatics and Political Engagement |
| Chen | Yi | Professor | MTSM | Data Mining, Databases |
| Curtmola | Reza | Professor | Computer Science | Security of Storage Systems, Security of Cloud Services, Software Security, Applied Cryptography, Security of Wireless Networks, Privacy in Computing |
| Ehrlich | Michael | Associate Professor | MTSM | Financial markets and institutions |
| Fjermestad | Jerry L. | Professor | MTSM | Management information systems |
| Fox | Wayne G. | Professor of Practice | MTSM | Corporate finance |
| Geller | James | Professor | Computer Science | Biomedical Informatics |
| Guo | Wenge | Associate Professor | Mathematical Sciences | Large-scale multiple testing, high dimensional data analysis, machine learning, bioinformatics, biostatistics |
| Halper | Michael | Professor | Informatics | Ontologies, Object-Oriented and Conceptual Modeling, Part-Whole Modeling, Modeling of Non-traditional Data, Mobile App Development, Computer Music. |
| Hoover | Amy | Assistant Professor | Informatics | AI in gaming and music |

| | | | | |
|---|---|---|---|---|
| Loh | Ji Meng | Associate Professor | Mathematical Sciences | Spatial Statistics |
| Kudyba | Stephan P. | Associate Professor | MTSM | Strategic Analytics, Healthcare Informatics, New technologies and operational transformation, Knowledge Management |
| Neamtiu | Iulian | Associate Professor | Computer Science | Software engineering, software aspects of machine and deep learning |
| Oria | Vincent | Professor | Computer Science | Multimedia databases, Spatio-Temporal Databases and Recommender Systems. |
| Perl | Yehoshua | Professor | Computer Science | Biomedical Informatics |
| Phan | Hai Nhat | Assistant Professor | Informatics | Adversarial attacks in AI models |
| Rohloff | Kurt | Associate Professor | Computer Science | Cybersecurity |
| Roshan | Usman | Associate Professor | Computer Science | Bioinformatics, machine learning |
| Shang | Zuofeng | Associate Professor | Mathematical Sciences | Nonparametric and semiparametric statistics, Bayesian methods, big data inference |
| Tang | Qiang | Assistant Professor | Computer Science | applied & theoretical cryptography, privacy and computer security. In particular: blockchain technology, post-Snowden cryptography, accountability, and information theoretic security. |
| Taylor (Fang) | Ming | Assistant Professor | MTSM | Earnings Management, Financial Reporting, Fraud and Regulation, Tax Avoidance Behavioral Accounting, Social Networks, Corporate Innovation, Emerging Markets |
| Taylor | Stephen | Assistant Professor | MTSM | application of mathematics, statistics, finance and data analysis/visualization. |
| Theodoratos | Dimitrios | Associate Professor | Computer Science | Databases, Data Management |
| Wang | Antai | Associate Professor | Mathematical Sciences | Biostatistics, Survival Analysis and High Dimensional Data Analysis |
| Wang | Guiling (Grace) | Professor | Computer Science | Data Mining |
| Wang | Jason | Professor | Computer Science | data mining and databases, data science and analytics, big data, machine learning, deep learning, computer vision, and services computing |
| Wei | Zhi | Professor | Computer Science | Bioinformatics, statistical inference |
| Wu | Chase | Professor | Computer Science | Big Data, machine learning, visualization |
| Wu | Yi-Fang (Brook) | Associate Professor | Informatics | text mining, information extraction, knowledge representation, information retrieval, and fake news early detection. |
| Yan | Zhipeng (Alan) | Associate Professor | MTSM | Empirical asset pricing and corporate finance |
| Yu | Dantong | Associate Professor | MTSM | Data Mining, high performance big data analysis |

All the faculty listed in the table are scholarly active, publish in high quality conferences and journals, and teach/advise in YWCC, MTSM, and Mathematics.

### VI.C. Computing Facilities, Equipment, Software and Datasets

NJIT's libraries, computing facilities and labs housed in the YWCC are satisfactory for the purposes of the proposed B.S. in Data Science.

The Van Houten Library: it offers access to numerous databases and peer-reviewed journals.

Venturelink: This is a community hub located at NJIT for technology companies to benefit from knowledge resources, expert mentors and advisors, and use co-working space for meetings and seminars.

Data observatory: This project is in progress.

Software and datasets: programming languages include C, Python, R, Tableau, Java and Perl. All software are free to access for NJIT students. In addition, NJIT has completely adopted Canvas system for course delivery.

### VI.D. Student Clubs

Data Science Club

This is a student-run association that holds events, lectures, and community activities, for students interested in machine learning. In addition, the club promotes job placement in data science careers.

ACM - The Association for Computing Machinery

The Association for Computing Machinery is the world's largest educational and scientific computing society. The student chapter at NJIT aims to build a community of students interested in the computing sciences, as well as encourage growth in the academic and professional fields.

G-WiCS

NJIT Graduate Women in Computing Society (G-WiCS) promotes and supports NJIT's growing community of women studying in the fields of computing and technology.

WiCS

NJIT Women in Computing Society (WiCS) provides similar functions as G-WiCS but is targeted to undergraduate students.

### VI.E. Funding

### VI.F. Industry Group

The YWCC maintains strategic relationships with industry, including Facebook, Merck, Prudential, and UPS, which will be leveraged for the Data Science program. The industry group provides capstone projects and mentoring to NJIT undergraduate students, and recruits from the NJIT graduates.

## VII. Degree Requirements

The Data Science curriculum includes 21 required courses, 4 electives, and a two semester senior project-based course. All of the required courses are from the CS and MATH departments.

In addition to the above areas, all undergraduate degree-seeking NJIT students must satisfy the university's General Education Requirements (GER). The GER is comprised of five thematic areas of knowledge and skills that the university has deemed necessary in order to improve and deepen students' critical thinking beyond the scope of their degree program. The thematic areas and corresponding requisite credit hours are as follows:

- Liberal Arts Literacy (18 credit hours)
- Computational Literacy (3 credit hours)
- Social Science Literacy (3 credit hours)
- Scientific Literacy (7 credit hours)
- Quantitative Reasoning/Mathematics (6 credit hours)

The B.S. in Data Science degree requires a minimum of 120 credit hours. Students may transfer up to 60 credit hours from other programs.

## VIII. Core Contributors

The core contributors to this proposal include (in alphabetical order): David A. Bader, Cristian Borcea, Craig Gotsman, David Horntrop, Ji Meng, Zoi-Heleni "Eliza" Michalopoulou, Connor Watson, Qishi "Chase" Wu.

## IX. References

New Jersey Institute of Technology. (2019a). *Annual Institutional Profile Report*. Retrieved from https://www5.njit.edu/oie/content/institutional-profle/

New Jersey Institute of Technology. (2019b). *Building a Strong Foundation—NJIT 2025*. Retrieved from https://www.njit.edu/strategicplan/njit-2025