



Data Sciences

2021-2022 HANDBOOK



PennState
College of Engineering

**ELECTRICAL ENGINEERING
AND COMPUTER SCIENCE**

Table of Contents	
TABLE OF CONTENTS	2
<u>WELCOME</u>	2
<u>THE DATA SCIENCES MAJOR</u>	4
PROGRAM OBJECTIVES	5
<u>ADVISING AND PROCEDURES FOR MAJOR</u>	5
ENTRANCE TO THE MAJOR (ETM)	6
DEGREE AUDITS	6
CONCURRENT MAJORS	6
REGISTRATION AND SUGGESTED ACADEMIC PLAN	6
PREREQUISITE COURSES	7
TRANSFER CREDIT	7
COOPERATIVE EDUCATION PROGRAM AND INTERNSHIPS	7
HONORS PROGRAM	8
MINORS	8
WAIVERS, EXCEPTIONS, AND PETITIONS	8
ACADEMIC INTEGRITY	8
<u>GRADUATION REQUIREMENTS</u>	9
C-REQUIREMENTS	9
DATA SCIENCES/IST (19 CREDITS)	9
COMPUTER SCIENCE AND ENGINEERING (24 CREDITS)	10
STATISTICS (11 CREDITS)	10
TECHNICAL ELECTIVES (12 CREDITS)	10
COMMUNICATIONS (9 CREDITS)	11
QUANTIFICATION (14 CREDITS)	11
NATURAL SCIENCES (9 CREDITS)	11
OTHER GENERAL EDUCATION (21 CREDITS)	12
DEPARTMENT LIST (GENERAL ELECTIVE) GUIDELINES (6 CREDITS)	12
FIRST YEAR SEMINAR (1 CREDIT)	13
<u>SOURCES OF INFORMATION</u>	1

Welcome

This handbook has been prepared for your use as a guide for your studies and as a means of providing you with much of the information that you may need as you continue to work towards your degree. We hope that you read it carefully, and we invite your inquiries about any of the

questions or issues that are related to your program. The Academic Affairs staff in W209 Westgate is here to serve you.

Please watch for announcements of special courses, news related to scheduling or textbooks, and other special opportunities delivered via a departmental email list.

Again, welcome to Computer Science and Engineering. We wish you well in your studies and offer our services to assist you.

Tom La Porta
Director, School of Electrical Engineering and Computer Science

Chita Das
Head, Department of Computer Science and Engineering

John Hannan
Associate Head, Department of Computer Science and Engineering

Mark Mahon
Undergraduate Advisor, School of Electrical Engineering and Computer Science

Alisha Simon
Academic Advisor, Department of Computer Science and Engineering

The Data Sciences Major

The Department of Computer Science and Engineering was created in 1993 with the merger of the Computer Engineering Program and the Computer Science Department. The department offers B.S. degrees in both computer science (CMPSC) and computer engineering (CMPEN) through the College of Engineering. It also offers the Computational Option of the inter-college Data Sciences B.S. degree.

The data sciences degree is part of an inter-college initiative between the College of Information Sciences and Technology (IST), College of Engineering, and Eberly College of Science to meet the need of professionals who can make sense of big data. The program provides students with the technical fundamentals of data science, with a focus on developing the knowledge and skills needed to manage and analyze large-scale, unstructured data to address an expanding range of problems in industry, government, and academia. As a result, data sciences graduates will possess the core skills and problem-solving approaches to compete for leading-edge analytics positions across many different industry sectors.

Computational Data Sciences, offered only through the Department of Computer Science and Engineering, focuses on the computational foundations of data science, including the design, implementation and analysis of software that manages the volume, heterogeneity and dynamic characteristics of large data sets and that leverages the computational power of multicore hardware. Students in this option will take upper-level courses in computer science and related fields to develop the skills necessary to construct efficient solutions to computational problems involving large data sets.

The mission of our undergraduate program is to prepare our students for a wide range of careers as computational data scientists and related positions in the field of computing. Our curriculum covers fundamental programming techniques and skills, broad knowledge of data science foundations, mathematical foundations of computing, and advanced topics in computing with large data sets. This curriculum provides students with the skills needed to design, develop, evaluate and analyze software solutions to computational problems involving large data and prepares them to be leaders throughout their careers. This program is intended to produce data science professionals with a deep understanding of how to compute with large data and not merely technicians who can use off-the-shelf tools. Success requires a strong aptitude in mathematics.

Because of the close relationships to Computational Data Sciences, concurrent majors in Computational Data Sciences and Computer Engineering or Computational Data Science and Computer Science are not permitted.

Program Objectives

Within a few years after graduation, graduates of the Computational Data Science major should be able to:

1. Apply appropriate theory, practices, and tools of data science to the specification, design, implementation, maintenance and evaluation of software that analyzes and manipulates large data sets.
2. Work and communicate effectively on multi-disciplinary teams.
3. Engage in continuous professional development through work assignments, graduate school study, professional training programs and independent learning.

During the first two years, students heading towards the Computational Data Sciences major take many courses in common with other engineering majors, including courses in mathematics. In addition, students take several specialized courses in the major, such as programming fundamentals. From these courses, students gain experience constructing software and completing individual and group projects. During the second two years, students complete a series of courses in computation theory, software systems and computing for data sciences. Students also select from numerous electives. Throughout the four years, students develop communication skills, including a senior year course in which students examine the complete design process and participate in a series of oral and written experiences similar to those that would be seen in industry.

Advising and Procedures for Major

If you are a student at University Park who is intending to major in Computational Data Sciences but who has not yet officially entered the major, you can make an appointment through Starfish to meet with an adviser at the **Engineering Advising Center (EAC)**, 208 Hammond Building, 863-1033.

If you are a student who has officially entered the DTSCE major, you can find your assigned adviser's name and email address in LionPATH. **If you communicate via e-mail, always use your Penn State account**, not another account such as Gmail. For non-advising issues (questions about e-petition status, course controls, etc.) see one of the CSE staff in W209 Westgate during regular business hours.

Information about all majors at Penn State is listed in the *Undergraduate Bulletin* at <https://bulletins.psu.edu/undergraduate/>. The *Bulletin* is updated yearly, so make sure to refer to the version of the Bulletin for the semester that you began at Penn State. Clarifications to the *Bulletin* are noted in this handbook, so it should be used in conjunction with the *Bulletin*.

The final responsibility for selecting courses and meeting degree requirements is yours. The role of your adviser is to suggest, recommend, and remind you of the requirements of the major and rules of the University.

Do not rely on LionPATH to correctly categorize your courses. LionPATH only understands the degree requirements as specified in the Bulletin and will therefore occasionally place courses into

an inappropriate category. This handbook contains additional restrictions and explanations. To ensure that you meet degree requirements, you must have your degree audit reviewed by your advisor periodically. You should submit petitions to correct any mis-categorization by LionPATH. Failure to do so may result in delaying your graduation until degree requirements are met.

Because data sciences is such a rapidly changing field, adjustments in course content and/or course offerings should be expected. It will be to your advantage to keep abreast of new course offerings, current course enhancements, and allowable course substitutions through regular contact with your adviser and the department office.

Entrance to the Major (ETM)

For the entrance-to-major requirements, see the “How to Get in” section of the Data Sciences major page in the University Bulletin for the year that you began at Penn State: <https://bulletins.psu.edu/undergraduate/archive/>

Degree Audits

Instructions on how to run a degree audit are available at https://tutorials.lionpath.psu.edu/public/S_RunningDegreeAudit/. You are responsible for periodically checking your degree audit on LionPATH to verify that the courses you have taken and plan to take will satisfy your degree requirements and that you are on track to complete your degree when you expect to complete it. You are encouraged to meet with an advisor to review your degree audit to verify this information. **Do not rely on LionPATH to place your courses into their appropriate categories.** Do not use a what-if report to check your degree requirements once you are in your major – what-if reports are for students who have not entered their major yet and may not show the correct set of requirements for students who have entered their major.

Concurrent Majors

Concurrent majors will not be allowed in Computational Data Sciences and Computer Engineering or Computational Data Sciences and Computer Science, although it is possible to obtain a concurrent major with another non-enrollment-controlled program.

Registration and Suggested Academic Plan

At least several weeks before it is time to register for the next semester’s courses, refer to this handbook and consult with your adviser to determine an appropriate set of courses. It is very important to schedule on your assigned scheduling date– **courses fill up quickly, and if you delay for even a few days, you may not be able to get into recommended courses for the next semester!**

For a copy of the suggested academic plan for the major, refer to the Computational Data Sciences major page in the University Bulletin for the year that you began at Penn State: <https://bulletins.psu.edu/undergraduate/archive/>. Re-ordering your course schedule from the suggested plan will not necessarily delay graduation. The key to completing 125 credits over 4 years is to average approximately 16 credits per semester. Though many students do maintain this pace, it is not unusual for students to take lighter loads some semesters and to delay graduation or to take summer classes. Some electives are not offered every semester and most third and fourth year classes are not offered in the summer, so please be careful in your scheduling.

Please realize that although all the courses listed on the plan are required for the degree, they need not be taken during the semesters shown in the charts. You may enroll in courses earlier than the semester that they are listed on the academic plan as long as you meet any prerequisites and controls on the courses. Classes that are not a prerequisite to any other courses can usually be taken in a later semester than they are listed on the plan without causing course sequencing issues. You should not wait until your last semester to take C-required courses.

You should be sure to check course prerequisites before you deviate from the suggested schedule. Care should be exercised to be sure core courses are taken in the proper sequence and in a timeframe allowing you to meet entrance to major requirements. In particular, students should proceed through the Math 140 > Math 141 > Math 230 > Stat 414 > Stat 415 > CMPSC 448 > DS 340W sequence by taking the next course in the sequence each semester.

Prerequisite Courses

If a CMPEN or CMPSC course has prerequisites, you must complete the prerequisite course before taking the successor course. For most courses an appropriate grade is a grade of D or higher. If the prerequisite course is a “Prescribed C or better” course and you receive a D you may register for the next course but you still must retake the prerequisite course. Waiving of prerequisites is rarely approved and requires approval by both the course instructor and the associate department head. **If you schedule a CMPEN or CMPSC course for which you have not satisfied the prerequisites, you will be removed from the course near the start of the semester even though LionPATH did not prevent you from scheduling the course – prerequisite checking for CMPEN and CMPSC classes is done manually by the CSE Department.**

Stat, Math, and DS subject courses have prerequisite checking at the time of enrollment. LionPATH will not allow you to enroll for these courses unless you meet the prerequisites.

Transfer Credit

In addition to taking courses at any Penn State campus, you may be able to earn credit by transferring credits from another school. Before taking a course at another university, use the transfer credit tool in LionPATH and check with your adviser to be sure the course will transfer usefully.

If you hope to use a transfer course that you have not taken yet to replace a CMPSC, CMPEN, or DS course, email a copy of the course syllabus to Prof. Hannan (jjh9@psu.edu) before enrolling in the course to check whether it will be allowed to count for that requirement. If Prof. Hannan approves the use of the transfer credit, you will then need to fill out a petition at <https://coursesub.psu.edu/> to have the transfer credit count correctly on your degree audit.

Note that CMPSC 465, DS 340W and DS 440 must be taken at Penn State.

Cooperative Education Program and Internships

The cooperative education program provides work experience by alternating periods of academic study and full-time employment in industry or government. The program typically starts at the beginning of the junior year and consists of three rotations, providing a cumulative work experience of one year.

If you have interest in the co-op program, you should obtain advising no later than your fourth semester from the designated co-op adviser, who will help you plan work and study schedules. You may earn up to 3 credits toward graduation in the Department List requirements.

If you prefer less of a time commitment, you can pursue one or more summer internships. You earn 1 credit per internship (maximum of 2 credits total) toward graduation in the Department List requirements.

Instructions for registering for co-op and internship credit can be found at <https://career.engr.psu.edu/students/undergraduate/intern-coop/credit-options/courses.aspx>. If you are not a formal co-op or internship student, you may still take related summer jobs; however, you may not claim credits for jobs you arrange outside of the formal programs.

Honors Program

Students in the Schreyer Honors College (Atherton Hall, 863-2635) may earn honors in Computational Data Sciences by completing a thesis with a member of the CSE faculty. See an honors adviser if you are interested in finding out more. (The department office, W209 Westgate Building, can identify the honors advisers for you).

Minors

A minor is a specialization of at least 18 credits that supplements a major. Some courses may concurrently meet the requirements of our major. Popular minors for students in our department include:

- 1) Entrepreneurship and Innovation
- 2) Engineering Leadership Development
- 3) Mathematics
- 4) Statistics

Waivers, Exceptions, and Petitions

In order to graduate from Penn State, all requirements on your degree audit must be marked “satisfied.” If you hope to use a course in a way that it does not automatically fill in on your audit, you must submit a petition at <https://coursesub.psu.edu/> so that it can be determined whether the substitution you are requesting will be permitted and, if so, your degree audit can be updated. Be sure to submit course substitution petitions prior to taking courses and prior to the semester in which you plan to graduate so that you have time to make adjustments if your petition is not approved.

Note that petitions that require College level approval (exceptions/waivers to College & University requirements, such as general education requirements) must be submitted BEFORE the semester in which you plan to graduate.

Academic Integrity

Recognizing not only the value of integrity in the academic environment, but also its value for the practicing data scientist and for society at large, we in the department urge you to act as a responsible professional while you are a student. Academic integrity is defined as follows in Faculty Senate rule 49-20:

“Academic integrity is the pursuit of scholarly activity free from fraud and deception and is an educational objective of this institution. Academic dishonesty includes, but is not limited to, cheating, plagiarizing, fabricating of information or citations, facilitating acts of academic dishonesty by others, having unauthorized possession of examinations, submitting work of another person or work previously used without informing the instructor, or tampering with the academic work of other students.”

The EECS School maintains a specific [Academic Integrity Statement](https://www.eecs.psu.edu/students/resources/EECS-CSE-Academic-Integrity.aspx) at <https://www.eecs.psu.edu/students/resources/EECS-CSE-Academic-Integrity.aspx> related to programming courses.

It is commonly accepted that people learn better if they can interact, discuss, and assist each other in solving problems and understanding concepts. Yet persons submitting identical homework papers overstep the bounds of beneficial interaction. The specific limits of acceptable collaboration will be spelled out by the instructor in each course in the course syllabus. The specifics may vary from course to course, but you are always responsible for keeping your work private and inaccessible to others. Do not, for any reason, show another student a part of your code or write code for another student. Do not put your code online in any location that might be publicly accessible. Any collaboration that exceeds these guidelines or the instructor’s guidelines will be considered cheating. Clearly, professionals share ideas, but they should not use another’s work without clear acknowledgement of who did the work. Academic dishonesty in any form is not condoned or tolerated.

GRADUATION REQUIREMENTS

To graduate from the University, every student must:

- (1) Complete the course requirements for his or her major (including earning Cs or higher in all C-required courses) and
- (2) Earn at least a 2.0 cumulative grade-point average for all courses taken at the University

C-requirements

Some courses for the major require a grade of C or higher to graduate. If you have officially entered your major, the easiest way to see which classes require a grade of C or higher is to check your degree audit. On your degree audit, some sections are labeled as requiring a C or higher, and the classes used toward those sections are the ones that require a grade of C or above. If a class is being used in a section of the degree audit that does not have a label about a C or higher being required, a D or above will fulfill that requirement.

If you have not declared your major yet, you can use the Bulletin to see the classes that are C-required for you. The academic plan, with C-required classes labeled, is available on the Data Sciences major page in the University Bulletin for the year that you began at Penn State: <https://bulletins.psu.edu/undergraduate/archive/>.

Data Sciences/IST (19 credits)

- DS 200 (4) – Introduction to Data Sciences **or** STAT 200 (4)

- DS 220 (3) – Data Management for Data Sciences (CMPSC 121 or 131)
- DS 300 (3) – Data Privacy and Security (DS 220)
- DS 340W (3) – Applied Data Sciences (DS 300, CMPSC 448)
- DS 410 (3) – Data Analytics at Scale ((CMPSC 122 or CMPSC 132) and DS 220)
- DS 440 (3) – Data Sciences Capstone (DS 220)

Computer Science and Engineering (24 credits)

- CMPSC 121 GQ (3) – Introduction to Programming Techniques (MATH 110 or MATH 140 concurrently or as a prerequisite) OR
CMPSC 131 (3) – Programming and Computation I Fundamentals (MATH 110 or MATH 140 concurrently or as a prerequisite)
- CMPSC 122 (3) – Intermediate Programming (CMPSC 121) OR
CMPSC 132 (3) - Programming and Computation II Data Structures (CMPSC 121 or CMPSC 131)
- CMPSC 221 (3) - Object Oriented Programming with Web-Based Applications (CMPSC 122 or CMPSC 132)
- CMPSC 360 (3) – Discrete Mathematics for Computer Science (Concurrent: CMPSC 122 or 132)
- CMPSC 442 (3) – Artificial Intelligence (CMPSC 221; Concurrent: CMPSC 465)
- CMPSC 448 (3) – Machine Learning (STAT 415, CMPSC 122 or 132)
- CMPSC 455 (3) – Introduction to Numerical Analysis I (MATH 220; MATH 230 or MATH 231; and 3 credits of programming)
- CMPSC 465 (3) – Data Structures and Algorithms (CMPSC 360 or MATH 311W)

Statistics (11 credits)

- STAT 184 (2) - Introduction to R (MATH 21)
- STAT 380 (3) – Data Science through Statistical Reasoning and Computation (STAT 184, STAT 200)
- STAT 414 (3) – Introduction to Probability Theory (MATH 230)
- STAT 415 (3) – Introduction to Mathematical Statistics (STAT 414)

Technical Electives (12 credits)

Select 6 credits from **List A**:

- CMPSC 450 (3) – Concurrent Scientific Programming (CMPSC 121, CMPSC 131, CMPSC 201 or CMPSC 202; MATH 220; MATH 230 or MATH 231)
- CMPEN 454 (3) – Fundamentals of Computer Vision (MATH 230 or MATH 231; CMPSC 121 or CMPSC 201)
- CMPSC 456 (3) – Introduction to Numerical Analysis II (CMPSC 455)

Note: Some courses are NOT offered every semester or even every year.

Select 6 credits from **List B:**

- CMPSC 431W (3) – Database Management Systems (CMPSC 221; ENGL 202C)
- EE 456 (3) – Introduction to Neural Networks (CMPSC 201 or CMPSC 202; MATH 220)
- IST 441 (3) - Information Retrieval and Organization (IST 210, IST 240)
- STAT 416 (3) - Stochastic Modeling (STAT 318 or STAT 414; MATH 230)
- STAT 440 (3) - Computational Statistics (STAT 200, STAT 415, MATH 220)

Note: Some courses are NOT offered every semester or even every year.

Communications (9 credits)

- ENGL 15 GWS (3) – Rhetoric and Composition (ENGL 30 or ESL 15 may be substituted)
- ENGL 202C GWS (3) – Technical Writing
- CAS 100 A/B (3) – Effective Speech

ENGL/CAS 137 & 138 may substitute for ENGL 15 and CAS 100 A/B

Quantification (14 credits)

- MATH 140 GQ (4) – Calculus with Analytic Geometry I
- MATH 141 GQ (4) – Calculus with Analytic Geometry II
- MATH 220 GQ (2) – Matrices
- MATH 230 (4) – Calculus and Vector Analysis (combination of MATH 231 (2) and MATH 232 (2) may be substituted)

Natural Sciences (9 credits)

Nine credits of Natural Science (GN) are required. Any GN courses except the following may be used: ASTRO 1, 6, 7N, 10, 11, 120, 140; all BISC courses; All CHEM below CHEM 110 (except 3 credits of CHEM 106 can be used); PHYS 250, 251, all PHYS below PHYS 211, GEOSC 20.

Other General Education (21 credits)

The Health and Wellness (GHW) requirement can be met by taking one 3-credit course or various credit combinations, most frequently two 1.5 credit courses (which can be taken in different semesters). A student who completes an ROTC program may use 3 credits of ROTC to satisfy the GHW requirement.

Details for the remaining General Education requirements can be found in the Baccalaureate Degree General Education Requirements at <https://bulletins.psu.edu/undergraduate/general-education/baccalaureate-degree-general-education-program/>.

Department List (General Elective) Guidelines (6 credits)

Choose enough credits to bring the total number of credits up to at least 126. If your US/IL course was not an Arts, Humanities, Social or Behavioral Sciences course, it may be counted in this list. These are sometimes called approved free electives or general electives, but the following restrictions apply:

- no courses not satisfying minimum requirements for baccalaureate degree program (see course descriptions in University *Bulletin*)
- no courses described as intended for non-science or non-technical majors in course description in University *Bulletin* (**You may take non-technical courses** but look at the *Bulletin* to be sure the description doesn't say "for non-science majors only").
- no courses similar or remedial to a required course or course already taken (when in doubt, check with your advisor before scheduling the course). For example, you may not include 2 credits of MATH 140A or 2 credits of CHEM 106.
- not ENGL 4, 5, or any other remedial English
- no more than 3 credits of ROTC
- no more than 6 credits of music performance courses
- no more than 3 additional credits of physical education
- no more than 3 credits of Cooperative Education
- no more than 2 credits of Engineering Internship
- no more than 3 credits of CHEM 106
- none of the following:
 - Astronomy (ASTRO) 1, 6, 7N, 10, 11, 120, 140
 - Biological Science (BISC) 1, 2, 3, 4
 - Chemistry (CHEM) 1, 3, 108, 101

- Computer Science (CMPSC) 100, 101, 200, 201, 203
- Cybersecurity Analytics and Operations (CYBER) 100
- Earth and Mineral Sciences (EMSC) 150
- English as a Second Language (ESL) 4
- Information Science & Technology (IST) 140, 220, 230, 240, 242, 261, 311, 361
- Language and Literacy Education (LLED) 5, 10
- Mathematics (MATH) 200, 201, MATH below 140
- Philosophy (PHIL) 12
- Physical Science (PHSC) 7
- Physics (PHYS) 250, 251, PHYS below 211
- Science, Technology, and Society (S T S) 150
- Speech Communication (CAS) 126, 283
- Statistics (STAT or MATH) below 318
- STAT (MATH) 318, STAT (MATH) 319, STAT 401, STAT (MATH) 414, STAT (MATH) 415, STAT (MATH) 418

Because most classes at Penn State (that you are not already using for another degree requirement) can count toward department list, there isn't a list of all classes that can be used – it would be very long. Rather, refer to the list above of types of courses that are not eligible for this requirement.

First Year Seminar (1 credit)

Most DTSCE majors take a 1-credit first-year seminar in either their first or second semester at Penn State. If you did not take a first-year seminar that was a separate 1-credit course, you must make up the first-year seminar credit by taking an extra credit of coursework that is acceptable for the department list requirement. Then, submit a petition at <https://coursesub.psu.edu/> requesting to use the extra credit of department list coursework to fulfill the first-year seminar requirement on your degree audit.

Sources of Information

This *Handbook* provides program information specifically for the undergraduate computational data sciences major. It should be used as a supplement to the *College of Engineering Undergraduate Programs Guide* that is available online. The information in this *Handbook* pertains to students who began at Penn State in Summer 2021, Fall 2021, or Spring 2022 semesters. Students who began at Penn State in an earlier year should refer to the appropriate earlier version of the Handbook. Students in pre-major (ENGR) status may use this Handbook as a reference for scheduling; however, your official degree requirements will be established when you enter the major. For information about the computer engineering degree, refer to the *Computer Engineering Undergraduate Handbook*. For information about the computer science degree, refer to the *Computer Science Undergraduate Handbook*. All of these documents are available in the department office, W209 Westgate Building and online at <http://eecs.psu.edu/students/undergraduate/Majors-Minors-Certificates.aspx>. (If you are at a campus other than University Park, you should contact the College of Engineering representative at your location).

Although this *Handbook* lists all requirements for the data sciences major, only those specific to data sciences are described in detail. Other general College and University requirements are discussed only briefly with references to more comprehensive supporting documents. Hard copies of these documents can be obtained from a Dean's office or local bookstore. Many are available on-line. A list of useful web resources is provided below. For easy reference, resource names are printed in bold throughout the *Handbook*.

School of EECS – <http://eecs.psu.edu>

Engineering Advising Center – <https://advising.engr.psu.edu>

Academic Advising Portal – <http://advising.psu.edu>

Bulletin of Baccalaureate Degree Programs – <http://bulletins.psu.edu/undergrad>

University Faculty Senate – <http://www.senate.psu.edu/policies/>

Student Affairs – <https://studentaffairs.psu.edu>

General Education – <https://gened.psu.edu>

LionPATH – <http://launch.LionPATH.psu.edu>

Association of Women in Computing - <http://www.awc.cse.psu.edu>

Association for Computing Machinery Student Chapter – <http://acm.psu.edu>