

Data Science Major

Program Requirements

- **Total units required:** 51-54

Required Courses and Practicum

- 12 core requirement courses (36 credits)
- Four elective courses (12 credits)
- One course in Ethics and Professional Responsibility (3 credits)
- Practicum requirement (3 credits if completed via independent study or a project-focused course)

The College of Arts & Sciences and McKelvey School of Engineering developed a new major that efficiently captures the intersection of mathematics and statistics with computer science for data science. The Bachelor of Arts in Data Science (BADs) will give students the formal foundation needed to understand the applicability and consequences of the various approaches to analyzing data with a focus on statistical modeling and machine learning.

McKelvey Engineering students who declare this major must fulfill the core course requirements listed below and all other requirements for the Applied Science degree in the McKelvey School of Engineering. They must also complete Engr 310 Technical Writing and 8 units of courses designated as NSM (Natural Sciences & Math) from Anthropology (L48 Anthro), Biology and Biomedical Sciences (L41 Biol), Chemistry (L07 Chem), Earth, Environmental, and Planetary Sciences (L19 EEPS), Physics (L31 Physics), or Environmental Studies (L82 EnSt).

Arts & Sciences students who declare this major must fulfill the distribution requirements and all other requirements for an AB degree in addition to the specific requirements listed below.

Data Science Core Requirements (CR)

Code	Title	Units
Math 131	Calculus I	3
Math 132	Calculus II	3
Math 233	Calculus III	3
Math 309	Matrix Algebra	3
SDS 3211	Statistics for Data Science I	3
SDS 4211	Statistics for Data Science II	3
SDS 439	Linear Statistical Models	3
CSE 131	Introduction to Computer Science	3
CSE 217A	Introduction to Data Science	3
CSE 247	Data Structures and Algorithms	3
CSE 314A	Data Manipulation and Management	3
CSE 417T or SDS 460	Introduction to Machine Learning Multivariate Statistical Analysis	3
Total Units		36

Notes:

1. Each of these core courses must be passed with a grade of C- or better.
2. AP credit can be applied for Math 131 Calculus I, Math 132 Calculus II, and Math 233 Calculus III. Students who have completed Math 203 Honors Mathematics I and Math 204 Honors Mathematics II will have this requirement waived.
3. CSE 131 Introduction to Computer Science may be waived after consultation with the director of undergraduate studies of the Department of Computer Science and Engineering.

Data Science Technical Electives

Four courses can be chosen from the list of approved electives given below, with the following caveats:

- At least one course from Statistics and Data Science (at the 400 level or above)
- At least one course from CSE (ending in S, T, M, or A)
- At most one course at the 200 level

List of Approved Data Science Technical Electives

Computer Science and Engineering

Code	Title	Units
CSE 237S	Programming Tools and Techniques	3
CSE 256A	Introduction to Human-Centered Design	3
CSE 311A	Introduction to Intelligent Agents Using Science Fiction	3
CSE 347	Analysis of Algorithms	3
CSE 359A	Signals, Data and Equity (Cannot be double-counted in EPR)	3
CSE 411A	AI and Society (Cannot be double-counted in EPR)	3
CSE 412A	Introduction to Artificial Intelligence	3
CSE 416A	Data Science for Complex Networks	3
CSE 417T	Introduction to Machine Learning (Cannot be double-counted in CR)	3
CSE 427S	Cloud Computing with Big Data Applications	3
CSE 435S	Database Management Systems	3
CSE 457A	Introduction to Visualization	3
CSE 514A	Data Mining	3
CSE 515T	Bayesian Methods in Machine Learning	3
CSE 517A	Machine Learning	3
CSE 518A	Human-in-the-Loop Computation	3
CSE 534A	Large-Scale Optimization for Data Science	3
CSE 543T	Algorithms for Nonlinear Optimization	3
CSE 559A	Computer Vision	3

Statistics and Data Science

Code	Title	Units
SDS 322	Biostatistics	3
SDS 420	Experimental Design	3
SDS 434	Survival Analysis	3
SDS 4392	Advanced Linear Statistical Models	3
SDS 459	Bayesian Statistics	3
SDS 460	Multivariate Statistical Analysis	3
SDS 461	Time Series Analysis	3
SDS 462	Mathematical Foundations of Big Data	3
SDS 475	Statistical Computation	3
SDS 494	Mathematical Statistics	3
SDS 495	Stochastic Processes	3
SDS 496	Topics in Statistics	3
SDS 5061	Theory of Statistics I	3
SDS 5062	Theory of Statistics II	3
SDS 5071	Advanced Linear Models I	3
SDS 5072	Advanced Linear Models II	3
SDS 5531	Advanced Statistical Computing I	3
SDS 5532	Advanced Statistical Computing II	3
SDS 5595	Topics in Statistics: Spatial Statistics	3
SDS 579	Topics in Statistics	3
SDS 586	Topics in Statistics	3

Mathematics

Code	Title	Units
Math 449	Numerical Applied Mathematics	3
Math 450	Topics in Applied Mathematics	3
Math 456	Topics in Financial Mathematics	3
Math 5047	Geometry/Topology III: Differential Geometry	3

Electrical and Systems Engineering

Code	Title	Units
ESE 4031	Optimization for Engineered Planning, Decisions and Operations	3
ESE 415	Optimization	3
ESE 427	Financial Mathematics	3

Energy, Environmental & Chemical Engineering

Code	Title	Units
EECE 202	Computational Modeling in Energy, Environmental and Chemical Engineering	3

Linguistics

Code	Title	Units
Ling 317	Introduction to Computational Linguistics	3

Ethics and Professional Responsibility Requirement (EPR)

- 3 units of courses from the following list:

List of EPR Course Options

Code	Title	Units
Engr 450F	Engineers in the Community (Engineering Ethics, Leadership and Conflict Management)	3
Engr 4501	Engineering Ethics and Sustainability	1
Engr 4502	Engineering Leadership and Team Building	1
Engr 4503	Conflict Management and Negotiation	1
Engr 520P	Presentation Skills for Scientists and Engineers	2
CSE 359A	Signals, Data and Equity (This course cannot be double-counted as an elective)	3
CSE 411A	AI and Society (Cannot be double-counted as an Elective)	3
MSB 512	Ethics in Biostatistics and Data Science	2

Practicum Requirement

- Students must complete an approved comprehensive data science project or experience for their practicum requirement. The practicum must be approved by the committee of data science faculty.
- The practicum experience should be completed during the next-to-last semester of study (i.e., the first semester of senior year). It is important that practicum plans be submitted for review prior to starting the project or course work to ensure the proposed work is sufficient for the objectives of the practicum. After-the-fact approvals are possible but not guaranteed.
- Appropriate practicum work is possible via Independent Study (CSE 400E or SDS 400 Undergraduate Independent Study) or via project-focused classes, including (but not limited to) CSE 437S Software Engineering Workshop and CSE 454A Software Engineering for External Clients. Students should contact course instructors in advance to identify the degree of agency the student will have over project selection and requirements.
- Internships related to data science can be used to fulfill the practicum. If the internship is paid, it cannot count for credit, but can satisfy the requirements.
- To initiate the approval process, majors through the McKelvey School of Engineering should contact the CSE undergraduate coordinator (cse.undergrad@email.wustl.edu) in the CSE department, and majors through Arts & Sciences should contact the Associate Chair in the Statistics and Data Science Department, José Figueroa-López.

Notes to All Majors in Statistics and Data Science

1. Upper-level courses have course numbers that begin with a "3" or higher (e.g., SDS 3200 Elementary to Intermediate Statistics and Data Analysis). Lower-level courses do not count toward upper-level data science requirements, even if they are cross-listed as an upper-level course in another department or program. For example, if SDS 2200 Elementary Probability and Statistics was cross-listed by another department as 3XXX, registering for that 3XXX course would not satisfy an upper-level data science requirement.
2. Certain approved substitutions are found on the Department of Statistics and Data Science website.

Course Substitutions

- ESE 326 can be taken in place SDS 3200. ESE 326 Probability and Statistics for Engineering and SDS 3200 Elementary to Intermediate Statistics and Data Analysis cannot both count toward a major or minor.
- Any course from another department that is cross-listed as a Mathematics L24 or Statistics and Data Science L87 course can count as an upper-level elective. Examples include Math 501C Theoretical Physics, L24 Math 440C, and L24 Math 403C. Such L24 or L87 courses always end with a "C."
- The following courses can count as upper-level electives:
 - Econ 4151 (this course can count as a statistics elective)
 - ESE 319, ESE 403

Distinctions in Data Science

Distinction

- For Distinction in Data Science, a student must take an additional two electives for a total of 6 electives.
- The student's GPA in the 6 electives must be at least 3.7. If the student takes additional courses that satisfy these requirements, the courses with the lowest grades may be omitted when calculating the GPA for this purpose.
- The electives need to be taken from the list of approved data science technical electives. Electives need to be passed with a grade of B or better.

High Distinction

- Complete all requirements for Distinction.
- Complete an honors thesis in either department (Statistics and Data Science or Computer Science & Engineering).

Highest Distinction

- Complete the requirements for High Distinction.
- Complete one of the two options described below:
 - **Qual Option:** Take either SDS 5061 Theory of Statistics I–SDS 5062 Theory of Statistics II or SDS 5071 Advanced Linear Models I–SDS 5072 Advanced Linear Models II, which corresponds to one of the qualifiers for the PhD in Statistics, earning a grade of B+ or better in each course.
 - **Course Option:** Complete three additional electives from the list of approved data science technical electives for a total of 12 courses. As with Distinction, the student's GPA in the 12 electives must be at least 3.7 with the student earning a grade of B+ or better in each course, and additional courses beyond 12 can be disregarded when calculating the GPA. At least four of these 12 courses must be from the Department of Statistics and Data Science, and at least four must be from the Department of Computer Science & Engineering.

Additional Information

Additional Requirements

- All Statistics and Data Science majors must take Math 131 Calculus I, Math 132 Calculus II, and Math 233 Calculus III. There are other ways to fulfill this requirement, including AP credit and Math 203 Honors Mathematics I or Math 204 Honors Mathematics II. Some students may obtain a waiver if they took similar courses before coming to Washington University.
- All required courses (both lower- and upper-level courses) must be completed with a letter grade of C- or better.
- Courses transferred from other accredited colleges and universities can be counted, with the following caveats, if they receive department approval:
 - a. Courses transferred from a two-year college (e.g., a community college) cannot be used to satisfy upper-level requirements.
 - b. At least half of the upper-level units required in a major or minor program must be fulfilled by courses taken in the major's home departments or by approved courses taken in Washington University-approved overseas study programs.
 - c. Courses from the School of Continuing & Professional Studies cannot be used to fulfill major requirements.
- No **upper-level** course used to satisfy a major requirement can be counted toward the requirements of any other major or minor (i.e., no double-counting of courses).
- At most, 3 units of independent study or research work can count toward the major requirements.
- A student cannot declare more than one major or minor in the department.

Latin Honors

At the time of graduation, the Department of Statistics and Data Science will recommend that a candidate receive Latin Honors (cum laude, magna cum laude, or summa cum laude) if that student has completed the department's requirements for High Distinction or Highest Distinction in Mathematics, including an Honors Thesis. The actual award of Latin Honors is managed by the College of Arts & Sciences.

The Honors Thesis

Arts & Sciences majors who want to be candidates for Latin Honors, High Distinction, or Highest Distinction must complete an honors thesis. Writing an honors thesis involves a considerable amount of independent work, reading, creating mathematics, writing a paper that meets acceptable professional standards, and making an oral presentation of the results.

Types of Projects

An honors thesis can take three forms:

1. A thesis that presents significant work by the student on one or more nontrivial statistics or probability problems.
2. A project in applied statistics that involves an in-depth analysis of a large data set. To do an honors thesis involving data analysis, it is usually necessary to have completed SDS 3200, SDS 493, and SDS 494 (or SDS 3211 Statistics for Data Science I and SDS 4211 Statistics for Data Science II) by the end of the junior year and to have the ability to work with statistical software such as SAS, R, or Python.
3. A substantial expository paper that follows independent study on an advanced topic under the guidance of a department faculty member. Such a report would involve the careful presentation of ideas and the synthesis of materials from several sources.

Process and Suggested Timeline

Junior Year, Spring Semester:

1. Talk with a faculty advisor about possible projects.
2. Complete the Honors Proposal Form and submit it to Dr. Figueroa-López.

Senior Year:

1. By the end of January, provide the advisor with a draft abstract and outline of the paper.
2. By the end of February, submit a rough draft, including an abstract, to the advisor.
3. The student and the advisor should agree on a date that the writing will be complete and on a date and time for the oral presentation in mid-March (the deadline is March 31).

Departmental Prizes

Each year, the department considers graduating majors for several departmental prizes and also awards a prize to a junior. Recipients are recognized at an annual awards ceremony in April where graduating majors each receive a certificate and a set of honors cords to be worn as part of the academic dress at Commencement. Awards are noted on the student's permanent university record.

Contact: José E. Figueroa-López
Email: sdsadvising@wustl.edu
Website: <https://sds.wustl.edu/>